Revisiting the theory of interferometric wide-field synthesis

J. Pety[1,2] and N. Rodríguez-Fernández[1]

[1] IRAM, 300 rue de la Piscine, 38406 Grenoble cedex, France.
  e-mail: pety@iram.fr, rodriguez@iram.fr
[2] LERMA, UMR 8112, CNRS and Observatoire de Paris, 61 avenue de l'Observatoire, 75014 Paris, France.

ABSTRACT

Context. After several generations of interferometers in radioastronomy, wide-field imaging at high angular resolution is today a major goal for trying to match optical wide-field performances.
Aims. All the radio-interferometric, wide-field imaging methods currently belong to the mosaicking family. Based on a 30 years old, original idea from Ekers & Rots, we aim at proposing an alternate formalism.
Methods. Starting from their ideal case, we successively evaluate the impact of the standard ingredients of interferometric imaging, *i.e.* the sampling function, the visibility gridding, the data weighting, and the processing of the short spacings either from single-dish antennas or from heterogeneous arrays. After a comparison with standard nonlinear mosaicking, we assess the compatibility of the proposed processing with 1) a method of dealing with the effect of celestial projection and 2) the elongation of the primary beam along the scanning direction when using the on-the-fly observing mode.
Results. The dirty image resulting from the proposed scheme can be expressed as a convolution of the sky brightness distribution with a set of wide-field dirty beams varying with the sky coordinates. The wide-field dirty beams are locally shift-invariant as they do not depend strongly on position on the sky: Their shapes vary on angular scales typically larger or equal to the primary beamwidth. A comparison with standard nonlinear mosaicking shows that both processing schemes are not mathematically equivalent, though they both recover the sky brightness. In particular, the weighting scheme is very different in both methods. Moreover, the proposed scheme naturally processes the short spacings from both single-dish antennas and heterogeneous arrays. Finally, the sky gridding of the measured visibilities, required by the proposed scheme, may potentially save large amounts of hard-disk space and cpu processing power over mosaicking when handling data sets acquired with the on-the-fly observing mode.
Conclusions. We propose to call this promising family of imaging methods *wide-field synthesis* because it explicitly synthesizes visibilities at a much finer spatial frequency resolution than the one set by the diameter of the interferometer antennas.

Key words. techniques: wide-field - techniques: interferometric - methods: data analysis - methods: image processing

1. Introduction

The instantaneous field of view of an interferometer is naturally limited by the primary beam size of the individual antennas. For the ALMA 12m-antennas, this field of view is $\sim 9''$ at 690 GHz and $\sim 27''$ at 230 GHz. The astrophysical sources in the (sub)-millimeter domain are often much larger than this, but still structured on much smaller angular scales. Interferometric wide-field techniques enable us to fully image these sources at high angular resolution. These techniques first require an observing mode that in one way or another scans the sky on spatial scales larger than the primary beam. The most common observing mode in use today, known as stop-and-go mosaicking, consists in repeatedly observing sky positions typically separated by half the primary beam size. The improvement of the tracking behavior of modern antennas now leads astronomers to consider on-the-fly observations, with the antennas slewing continuously across the sky. The improvements in correlator and receiver technologies are also leading to techniques that could potentially sample the antenna focal planes with multi-beam receivers instead of the single-pixel receivers installed on current interferometers.

Send offprint requests to: e-mail: pety@iram.fr

The ideal measurement equation of interferometric wide-field imaging is

$$V(u_p, \alpha_s) = \int_{\alpha_p} B(\alpha_p - \alpha_s)\, I(\alpha_p)\, \mathrm{e}^{-i2\pi\alpha_p u_p}\, \mathrm{d}\alpha_p, \qquad (1)$$

where V is the visibility function of 1) u_p (the spatial frequency with respect to the fixed phase center) and 2) α_s (the scanned sky angle), I is the sky brightness, and B the antenna power pattern or primary beam of an antenna of the interferometer (Thompson et al. 1986, chapter 2). For simplicity, 1) we assume that the primary beam is independent of azimuth and elevation, and 2) we use one-dimensional notation without loss of generality. We do not deal with polarimetry (see e.g. Hamaker et al. 1996; Sault et al. 1996a, 1999) because it is adds another level of complexity over our first goal here, *i.e.* wide-field considerations. Several aspects make Eq. 1 peculiar with respect to the ideal measurement equation for single-field observations. First, the visibility is a function not only of the uv spatial frequency (u_p) but also of the scanned sky coordinate (α_s). Second, Eq. 1 is a mix between a Fourier transform and a convolution equation. It can be regarded, for example, as the Fourier transform along the α_p dimension

of the function, $B(\alpha_p - \alpha_s) \, I(\alpha_p)$, of the (α_p, α_s) variables. But Eq. 1 can also be written as the convolution:

$$V(u_p, \alpha_s) = \int_{\alpha_p} \mathcal{B}(\alpha_s - \alpha_p) \, \mathcal{I}(\alpha_p, u_p) \, d\alpha_p, \tag{2}$$

where $\quad \mathcal{B}(\alpha_s - \alpha_p) \equiv B(\alpha_p - \alpha_s) \tag{3}$

and $\quad \mathcal{I}(\alpha_p, u_p) \equiv I(\alpha_p) \, e^{-i2\pi\alpha_p u_p}. \tag{4}$

For each u_p kept constant, $V(u_p, \alpha_s)$ is the convolution of $\mathcal{B}$ and $\mathcal{I}$. Indeed, $\mathcal{I}(\alpha_p, u_p = 0) = I(\alpha_p)$, so we derive

$$V(u_p = 0, \alpha_s) = \int_{\alpha_p} \mathcal{B}(\alpha_s - \alpha_p) \, I(\alpha_p) \, d\alpha_p, \tag{5}$$

i.e., the convolution equation for single-dish observations.

Ekers & Rots (1979) were the first to show that the measurement equation (Eq. 1) enables us to recover spatial frequencies of the sky brightness at a much finer uv resolution than the uv resolution set by the diameter of the interferometer antennas. Interestingly enough, the goal of Ekers & Rots (1979) was "just" to find a way to produce the missing short spacings of a multiplying interferometer. However, Cornwell (1988) realized that Ekers & Rots' scheme has a much stronger impact, because it explains why an interferometer is able to do wide-field imaging. Cornwell (1988) also demonstrated that on-the-fly scanning is not absolutely necessary to interferometric wide-field imaging. Indeed, the large-scale information can be retrieved in mosaics of single-field observations, provided that the sampling of the single fields follows the sky-plane Nyquist sampling theorem.

As a result, all the information about the sky brightness is coded in the visibility function. From a data-processing viewpoint, all the current radio-interferometric wide-field imaging methods (see, e.g., Gueth et al. 1995; Sault et al. 1996b; Cornwell et al. 1993; Bhatnagar & Cornwell 2004; Bhatnagar et al. 2008; Cotton & Uson 2008) belong to the mosaicking family[1] pioneered by Cornwell (1988). In this family, the processing starts with Fourier transforming $V(u_p, \alpha_s)$ along the u_p dimension (*i.e.* at constant α_s) to produce a set of single-field dirty images before linearly combining them and forming a wide-field dirty image. In this paper, we propose an alternate processing, which starts with a Fourier transform of $V(u_p, \alpha_s)$ along the α_s dimension (*i.e.* at constant u_p). We show how this explicitly synthesizes the spatial frequencies needed to do wide-field imaging, which are linearly combined to form a "wide-field uv plane", *i.e.*, one uv-plane containing all the spatial frequency information measured during the wide-field observation. Inverse Fourier transform will produce a dirty image, which can then be deconvolved using standard methods. The existence of two different ways to extract the wide-field information from the visibility function raises several questions: Are they equivalent? What are their relative merits?

We thus aim at revisiting the mathematical foundations of wide-field imaging and deconvolution. Sections 2 to 7 propose the new algorithm, which we call wide-field synthesis: Section 2 first defines the notations and it then lays out the

basic concepts used throughout the paper. Section 3 states the steps needed to go beyond the Ekers & Rots scheme and explores the consequences of incomplete sampling of both the uv and sky planes. Section 4 discusses the effects of gridding by convolution and regular resampling. Section 5 describes how to influence the dirty beam shapes and thus the deconvolution. Section 6 states how to introduce short spacings measured either from a single-dish antenna or from heterogeneous interferometers. Section 7 compares the proposed wide-field synthesis algorithm with standard non-linear mosaicking. Some detailed demonstrations are factored out in Appendix A to enable an easier reading of the main paper, while ensuring that interested readers can follow the demonstrations. Appendices B and C then explain how the wide-field synthesis algorithm can cope with non-ideal effects: Appendix B discusses how at least one standard way to cope with sky projection problems is compatible with the wide-field synthesis algorithm. Appendix C explores the consequences of using the on-the-fly observing mode. Finally, we assume good familiarity with single-field imaging in various places. We refer the reader to well-known references: e.g. chapter 6 of Thompson et al. (1986) and Sramek & Schwab (1989).

2. Notations and basic concepts

2.1. Notations

In this paper, we use the Bracewell (2000)'s notation to display the relationship between a function $I(\alpha)$ and its direct Fourier transform $\overline{I}(u)$, *i.e.*,

$$I(\alpha) \supset \overline{I}(u), \tag{6}$$

where (α, u) is the couple of Fourier conjugate variables. We also use the following sign conventions for the direct and inverse Fourier transforms

$$\overline{I}(u) \equiv \int_{\alpha} I(\alpha) \, e^{-i2\pi\alpha u} \, d\alpha \tag{7}$$

and

$$I(\alpha) \equiv \int_{u} \overline{I}(u) \, e^{+i2\pi u\alpha} \, du. \tag{8}$$

As V is a function of two independent quantities (u_p and α_s), the Fourier transform may be applied independently on each dimension, while the other dimension stays constant. Several additional conventions are used to express this. First, we introduce a specific notation to state that either the first or the second dimension stays constant:

$$V_{u_p}(\alpha_s) \equiv V(u_p = \text{constant}, \alpha_s), \tag{9}$$

and

$$V^{\alpha_s}(u_p) \equiv V(u_p, \alpha_s = \text{constant}). \tag{10}$$

Second, we use a bottom/top line to derive the notation of the Fourier transform along the first/second dimension from the notation of the original function. Third, on the Fourier transform sign (*i.e.* $\supset$), we explicitly state the dimension along which the Fourier transform is computed. For instance, if D is a function of (α_p, α_s), then the Fourier transform of D along the first dimension is expressed as

$$D(\alpha_p, \alpha_s) \underset{u_p}{\overset{\alpha_p}{\supset}} \underline{D}(u_p, \alpha_s), \tag{11}$$

---

[1] In the rest of this paper, stop-and-go mosaicking refer to the observing mode, while mosaicking alone refer to the imaging family.

Table 1. Definition of the symbols used to expose the wide-field synthesis formalism.

Symbol & Definition		Plane(s)[a]
α_s	Scanned angle	sky
u_s	Scanned spatial frequency	uv
α_p	Phased angle	sky
u_p	Phased spatial frequency	uv
I	Sky brightness	sky
B	Primary beam	sky
V	Visibility function	uv & sky
S	Sampling function	uv & sky
Δ	Set of single-field dirty beams	sky & sky
D	Set of wide-field dirty beams	sky & sky
Ω	Sky-plane weighting function	sky & sky
W	uv-plane weighting function ($\Omega \supset W$)	uv & uv
$\mathcal{G}$	Gridding function ($= g\,\gamma$)	uv & sky
g	uv-plane gridding function	uv
γ	Sky-plane gridding function	sky
I_{dirty}	Wide-field dirty image	sky

Notes. [a] Planes of definition of the associated symbols.

while the Fourier transform of D along the second dimension is expressed as

$$D(\alpha_p, \alpha_s) \overset{\alpha_s}{\underset{u_s}{\supset}} \overline{D}(\alpha_p, u_s). \qquad (12)$$

We also use a more compact notation when doing the Fourier transform on both dimensions simultaneously, *i.e.*,

$$D(\alpha_p, \alpha_s) \overset{(\alpha_p, \alpha_s)}{\underset{(u_p, u_s)}{\supset}} \overline{\overline{D}}(u_p, u_s). \qquad (13)$$

Finally, the convolution of two functions G and V is noted and defined as

$$\{G \star V\}(u) \equiv \int_v G(u - v)\, V(v)\, \mathrm{d}v. \qquad (14)$$

For reference, Table 1 summarizes the definitions of the symbols used most throughout the paper. With the one-dimensional notation used throughout the paper, the number of planes quoted directly gives the number of associated dimensions of the symbols. Generalization to images would require a doubling of the number of planes/dimensions. Table 2 defines the uv and angular scales that are relevant to wide-field interferometric imaging, and Fig. 1 sketches the different angular scales. Each angular scale (θ) is related to a uv scale (d) through $\theta = 1/d$, where θ and d are measured in radians and in units of λ (the wavelength of the observation). In the rest of the paper, we explicitly distinguish between $\theta_{\mathrm{prim}} \equiv 1/d_{\mathrm{prim}}$, the angular scale associated to the diameter of the interferometer antennas, and θ_{fwhm}, the full width at half maximum of the primary beam. The relation between θ_{prim} and θ_{fwhm} depends on the illumination of the receiver feed by the antenna optics. In radio astronomy, we typically have $\theta_{\mathrm{fwhm}} \sim 1.2\,\theta_{\mathrm{prim}}$ (see e.g. Goldsmith 1998, chapter 6). Finally, the notion of anti-aliasing scale (θ_{alias}) is introduced and discussed in Sect. 4.2.

2.2. Basic concepts

Figure 2 illustrates the principles underlying 1) the setup to get interferometric wide-field observations and 2) our

Table 2. Definition of the uv and sky scales relevant to wide-field interferometric imaging.

Symbol	Definition
$[\lambda, \mathrm{rad}]^a$	Conjugate uv and angular scale
$d_{\max}, \theta_{\mathrm{syn}}$	Maximum baseline length & Synthesized beam
$d_{\mathrm{prim}}, \theta_{\mathrm{prim}}$	Antenna diameter & Primary beamwidth
$d_{\mathrm{alias}}, \theta_{\mathrm{alias}}$	Minimum image size for tolerable aliasing
$d_{\mathrm{field}}, \theta_{\mathrm{field}}$	Targeted field of view
$d_{\mathrm{image}}, \theta_{\mathrm{image}}$	Final image size

Notes. [a] The chosen units (radians for θ and wavelength for d) imply that the conjugate scales are linked through $\theta = 1/d$, instead of the usual $\theta = \lambda/d$.
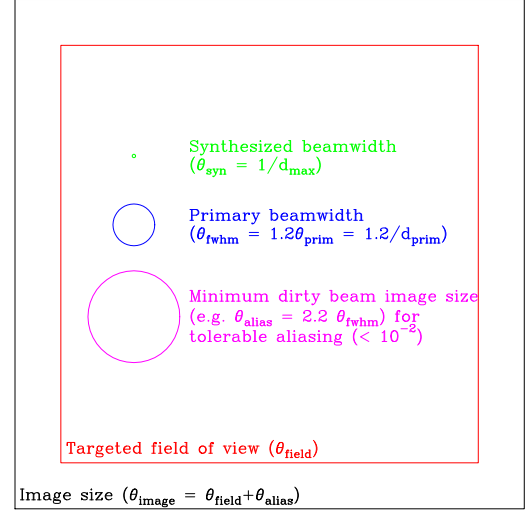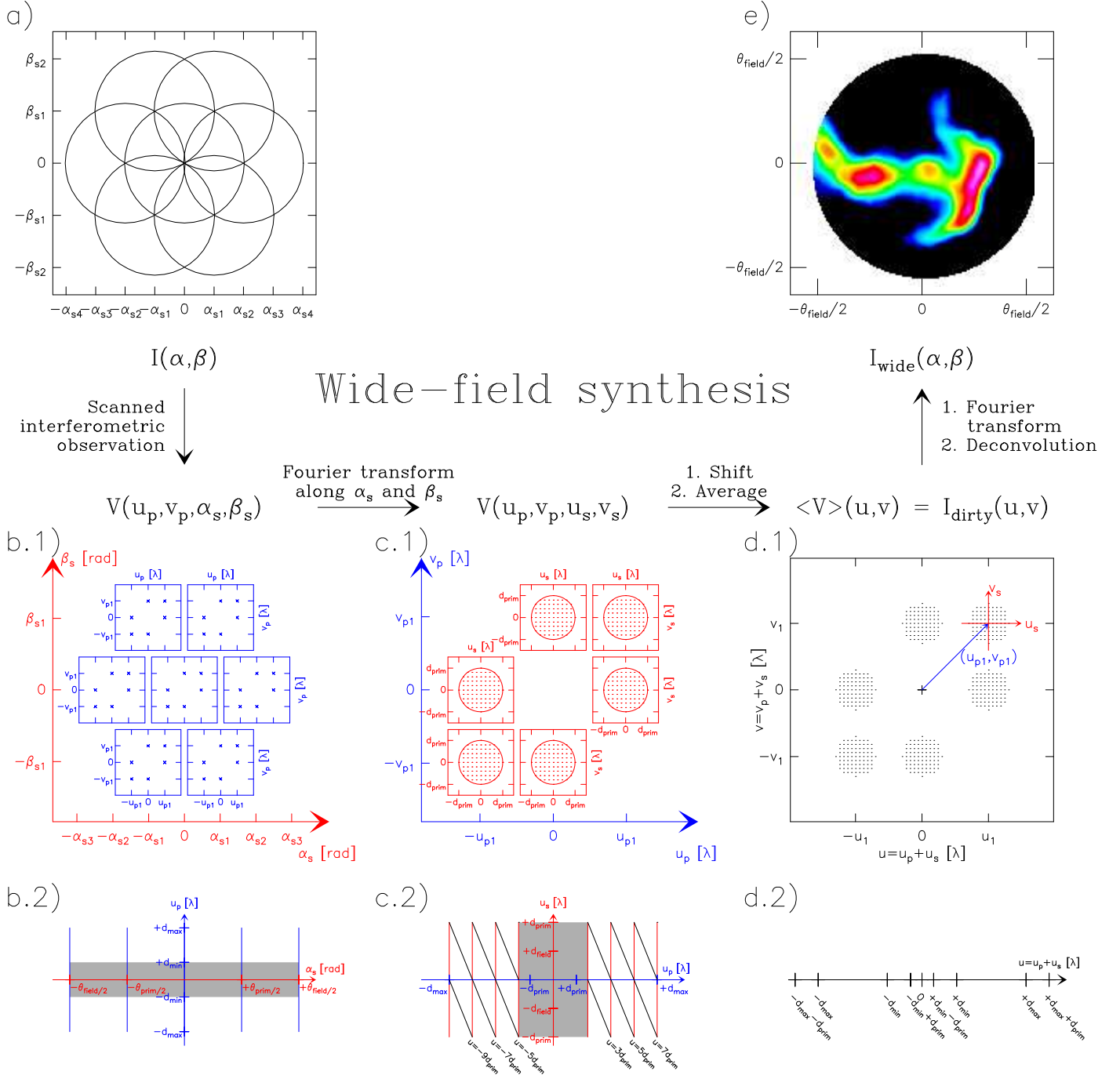

Fig. 1. Visualization of the different angular scales relevant to wide-field interferometric imaging. The notion of anti-aliasing scale (θ_{alias}) is introduced and discussed in Sect. 4.2.

proposition to process them. For simplicity, we display the minimum possible complexity without loss of generality. The top row displays the sky plane. The middle row represents the 4-dimensional measurement space at different stages of the processing. As it is difficult to display a 4-dimensional space on a sheet of paper, the bottom row shows 2-dimensional cuts of the measurement space at the same processing stages.

2.2.1. Observation setup and measurement space

Panel a) displays the sky region for which we aim for estimating the sky brigthness, $I(\alpha)$. The field of view of an interferometer observing in a given direction of the sky has a typical size set by the primary beam shape. In our example, this is illustrated by any of the circles whose diameter is θ_{prim}. As we aim at observing a wider field of view, e.g. θ_{field}, the interferometer needs to scan the targeted sky field. We assume that we scan through stop-and-go mosaicking, ending up with a 7-field mosaic.

After calibration, the output of the interferometer is a visibility function, $V(u_p, \alpha_s)$, whose relation to the sky brightness is given by the measurement equation (Eq. 1). Panel b.1) shows the measurement space as a mosaic of single-field uv planes: The uv plane coverage of each single-field observation is displayed as a blue sub-panel at the

a)

β_{s2}
β_{s1}
0
$-\beta_{s1}$
$-\beta_{s2}$

$-\alpha_{s4}$ $-\alpha_{s3}$ $-\alpha_{s2}$ $-\alpha_{s1}$ 0 α_{s1} α_{s2} α_{s3} α_{s4}

e)

$\theta_{field}/2$
0
$-\theta_{field}/2$

$-\theta_{field}/2$ 0 $\theta_{field}/2$

$I(\alpha,\beta)$

Wide−field synthesis

$I_{wide}(\alpha,\beta)$

Scanned interferometric observation

1. Fourier transform
2. Deconvolution

$V(u_p,v_p,\alpha_s,\beta_s)$

Fourier transform along α_s and β_s

$V(u_p,v_p,u_s,v_s)$

1. Shift
2. Average

$<V>(u,v) = I_{dirty}(u,v)$

b.1)

β_s [rad]
β_{s1}
0
$-\beta_{s1}$

u_p [λ]
v_{p1} 0 $-v_{p1}$

$-\alpha_{s3}$ $-\alpha_{s2}$ $-\alpha_{s1}$ 0 α_{s1} α_{s2} α_{s3} α_s [rad]

$-u_{p1}$ 0 u_{p1}

c.1)

v_p [λ]
v_{p1}
0
$-v_{p1}$

u_s [λ]
d_{prim} 0 $-d_{prim}$

$-u_{p1}$ 0 u_{p1}

$-d_{prim}$ 0 d_{prim}

u_p [λ]

d.1)

$v=v_p+v_s$ [λ]
v_1
0
$-v_1$

(u_{p1},v_{p1})

v_s
u_s

$-u_1$ 0 u_1
$u=u_p+u_s$ [λ]

b.2)

u_p [λ]
$+d_{max}$
$+d_{min}$
$-d_{min}$
$-d_{max}$

$-\theta_{field}/2$ $-\theta_{prim}/2$ $+\theta_{prim}/2$ $+\theta_{field}/2$ α_s [rad]

c.2)

u_s [λ]
$+d_{prim}$
$+d_{field}$
$-d_{field}$
$-d_{prim}$

$-d_{max}$ $-d_{prim}$ $+d_{prim}$ $+d_{max}$ u_p [λ]

$u=-9d_{prim}$ $u=-7d_{prim}$ $u=-5d_{prim}$ $u=3d_{prim}$ $u=5d_{prim}$ $u=7d_{prim}$

d.2)

$u=u_p+u_s$ [λ]

$-d_{max}-d_{prim}$ $-d_{max}$ $-d_{min}$ $-d_{min}+d_{prim}$ 0 $+d_{min}-d_{prim}$ $+d_{min}$ $+d_{max}$ $+d_{max}+d_{prim}$

Fig. 2. Illustration of the principles of wide-field synthesis, which enables us to image wide-field interferometric observations. The top row displays the sky plane. The middle row displays the 4-dimensional visibility space and the bottom row shows 2-dimensional cuts of this space at different stages of the processing. In panels b) to d), the scanned dimensions (α_s and u_s) are displayed in blue while the phased spatial scale dimensions (u_p) are displayed in red and the spatial scale dimensions (u) of the final wide-field uv plane are displayed in black. The grey zones of panels b.2) and c.2) show the regions of the visibility space without measurements (missing short-spacings). In detail, panel a) shows a possible scanning strategy of the sky to measure the unknown brightness distribution at high angular resolution: For simplicity it is here just a 7-field mosaic. Panel b.1) and b.2) sketch the space of measured visibilities: The uv plane at each of the 7 measured sky positions is displayed as a blue square box in panel b.1) and a blue vertical line in panel b.2). For simplicity, only 6 visibilities are plotted in panel b.1). Panels c.1) and c.2) sketch the space of synthesized visibilities after Fourier transform of the measured visibilities along the scanned coordinate (α_s): At each measured spatial frequency u_p (displayed on the blue axes) is associated one space of synthesized wide-field spatial frequencies displayed as one of the red squares in panel c.1) and the red vertical lines in panel c.2). The wide-field spatial scales are synthesized 1) on a grid whose cell size is related to the total field of view of the observation and 2) only inside circles whose radius is the primary diameter of the interferometer antennas. Panels d.1) and d.2) display the final, wide-field uv plane. This plane is built by application of the shift-and-average operator along the black lines on panel c.2), lines that display the region of constant u spatial frequency in the (u_p,u_s) space. Standard inverse Fourier transform and deconvolution methods then produce a wide-field distribution of sky brightnesses as shown in panel e).

sky position where it has been measured and which is featured by the red axes. We assume 1) that the interferometer has only 3 antennas and 2) that only a single integration is observed per sky position. This implies only 6 visibilities per single-field uv plane. In panel b.2), the uv planes at constant α_s are displayed as the blue vertical lines. The measured spatial frequencies belong to the $[-d_{\max}, -d_{\min}]$ and $[+d_{\min}, +d_{\max}]$ ranges, where $d_{\min}$ and $d_{\max}$ are respectively the shortest and longest measured baseline length. $d_{\min}$ is related to the minimum tolerable distance between two antennas to avoid collision. Here, we chose $d_{\min} \sim 1.5\, d_{\mathrm{prim}}$. The grey zone between $-d_{\min}$, and $+d_{\min}$ displays the missing short spacings.

2.2.2. Processing by explicit synthesis of the wide-field spatial frequencies

All the information about the sky brightness, $I(\alpha)$, is somehow coded in the visibility function, $V(u_p, \alpha_s)$. The high spatial frequencies (from $d_{\min}$ to $d_{\max}$) are clearly coded along the u_p dimension. The uncertainty relation between Fourier conjugate quantities also implies that the typical spatial frequency resolution along the u_p dimension is only d_{prim} because the field of view of a single pointing has a typical size of θ_{prim}. However, wide-field imaging implies measuring all the spatial frequencies with a finer resolution, $d_{\mathrm{field}} = 1/\theta_{\mathrm{field}}$. The missing information must then be hidden in the α_s dimension.

In Sect. 3, we show that Fourier transforming the measured visibilities along the α_s dimension (*i.e.* at constant u_p) can synthesize the missing spatial frequencies, because the α_s dimension is sampled from $-\theta_{\mathrm{field}}/2$ to $+\theta_{\mathrm{field}}/2$, implying a typical spatial-frequency resolution of the u_s dimension equal to d_{field}. Conversely, the α_s dimension is probed by the primary beams with a typical angular resolution of θ_{prim}, implying that the u_s spatial frequencies will only be synthesized inside the $[-d_{\mathrm{prim}}, +d_{\mathrm{prim}}]$ range. Panels c.1) and c.2) illustrate the effects of the Fourier transform of $V(u_p, u_s)$ along the α_s dimension, in 4 and 2 dimensions, respectively. The red subpanels or vertical lines display the u_s spatial frequencies around each constant u_p spatial frequency.

In panels d.1) and d.2) (*i.e.* after the Fourier transform along the α_s dimension), $\overline{V}(u_p, u_s)$ contains all the measured information about the sky brightness in a spatial frequency space. However, the information is ordered in a strange and redundant way. Indeed, we show that $\overline{V}(u_p, u_s)$ is linearly related to $\overline{I}(u_p + u_s)$. To first order, the information about a given spatial frequency u is stored in all the values of $\overline{V}(u_p, u_s)$ which verifies $u = u_p + u_s$ (black lines on panel c.2).

A shift operation will reorder the spatial scale information and averaging will compress the redundancy (illustrated by the halving of the number of the space dimensions). The use of a shift-and-average operator thus produces a final uv plane containing all the spatial scale information to image a wide field in an intuitive form. We thus call this space the *wide-field uv plane*. Panels d.1) and d.2) display this space, where the minimum relevant spatial frequency is related to the total field of view, while the maximum one is related to the interferometer resolution.

Sections 3 and 4 show that applying the shift-and-average operator to $\overline{V}$ produces the Fourier transform of a dirty image, which is a local convolution of the sky brightness by a slowly varying dirty beam. As a result, inverse Fourier transform of $\langle \overline{V} \rangle$ and deconvolution methods will produce a wide-field distribution of sky brightness as shown in panel e) at the top right of Fig. 2.

3. Beyond the Ekers & Rots scheme

In the real world, the visibility function is not only sampled, but this sampling is incomplete for two main reasons. 1) The instrument has a finite spatial resolution, and the scanning of the sky is limited, implying that the sampling in both planes has a finite support. 2) The uv coverage and the sky-scanning coverage can have holes caused either by intrinsic limitations (e.g. lack of short spacings or small number of baselines) or by acquisition problems (implying data flagging). The incomplete sampling makes the mathematics on the general case complex. We thus start with the ideal case where we assume that the visibility function is continuously sampled along the u_p and α_s dimension. We then look at the general case.

3.1. Ideal case: Infinite, continuous sampling

Starting from the measurement equation 1, Ekers & Rots (1979) first demonstrated (see Sect. A.1) that[2]

$$\forall (u_p, u_s), \quad \overline{V}_{u_p}(u_s) = \overline{B}(-u_s)\,\overline{I}(u_p + u_s). \tag{15}$$

For each constant u_p spatial frequency, the Fourier transform thus synthesizes a function, $\overline{V}_{u_p}(u_s)$, which is simply related to $\overline{I}(u_p + u_s)$, the Fourier components of the sky brightness around u_p. $\overline{V}(u_p, u_s)$ is only defined in the $[-d_{\mathrm{prim}}, +d_{\mathrm{prim}}]$ interval along the u_s dimension because $\overline{B}(-u_s)$ is itself only defined inside this interval, since $\overline{B}(-u_s)$ is the autocorrelation of the antenna illumination.

We search to derive a single estimate of the Fourier components $\overline{I}(u)$ of the sky brightness. Equation 15 indicates that the fraction $\overline{V}(u_p, u_s)/\overline{B}(-u_s)$ gives us an estimate of $\overline{I}(u)$ for each couple (u_p, u_s) that satisfies $u = u_p + u_s$. However, the information about $\overline{I}$ is strangely ordered. There are two possible ways to look at this ordering. 1) Starting from the measurement space, the Ekers & Rots scheme synthesizes frequencies around each u_p measure inside the interval $[u_p - d_{\mathrm{prim}}, u_p + d_{\mathrm{prim}}]$ at the d_{field} spatial frequency resolution. 2) Starting from our goal, we want to estimate $\overline{I}$ at a given spatial frequency u with a d_{field} spatial frequency resolution. We thus search for all the couples (u_p, u_s) satisfying $u = u_p + u_s$, which are displayed in panel c.2) of Fig. 2 as the diagonal black lines. It immediately results that 1) there are several estimates of $\overline{I}$ for each spatial frequency u and 2) the number of estimates varies with u. We can average them to get a better estimate of $\overline{I}(u)$.

---

[2] The convolution theorem, which states that the Fourier transform of the convolution of two functions is the product of the Fourier transform of both individual functions, is a special case for Eq. 15: It can be recovered by setting $u_p = 0$. Indeed, as already mentioned in the introduction, the ideal measurement equation 1 can be interpreted as a convolution with an additional phase term. By Fourier transforming along the α_s dimension, the convolution translates into a product of Fourier transforms $\overline{B}$ and $\overline{I}$, while the phase term translates into a shift of coordinates: $u_p + u_s$.

This last viewpoint thus suggests averaging in the (u_p, u_s) space along linepaths defined by $u = u_p + u_s$. Such an operator can mathematically be defined as

$$\langle F \rangle (u) \equiv \iint_{u_p u_s} \delta[u - (u_p + u_s)] \, W(u_p, u_s) \, F(u_p, u_s) \, \mathrm{d}u_p \mathrm{d}u_s, \quad (16)$$

where F is the function to be averaged and W is a normalized weighting function. Using the properties of the Dirac function, we can reduce the double integral to

$$\langle F \rangle (u) = \int_{u_p} W(u_p, u - u_p) \, F(u_p, u - u_p) \, \mathrm{d}u_p. \quad (17)$$

In this equation, we easily recognize a *shift-and-average operator*. The normalized weighting function plays a critical role in the following formalism, and we propose clever ways to define W in Sect. 5. In the ideal case studied here, W can be defined as

$$W(u_p, u_s) \equiv 1/(2\sqrt{2} \, d_{\mathrm{prim}}) \text{ for } u_s[-d_{\mathrm{prim}}, +d_{\mathrm{prim}}],$$
$$W(u_p, u_s) \equiv 0 \text{ for other values of } u_s.$$

In other words, we have just normalized the integral by the constant length $(2\sqrt{2} \, d_{\mathrm{prim}})$ of the averaging linepath.

3.1.1. Wide-field dirty image, dirty beam and image-plane measurement equation

Section 3.2 shows that the incomplete sky and uv sampling forbid us to apply the shift-and-average operator to the $\overline{V}(u_p, u_s)/\overline{B}(-u_s)$ function. To guide us in this general case, we thus explore the consequences of applying this operator to $\overline{V}$ in the ideal case. It is easy to demonstrate that the result is the Fourier transform of a dirty image, *i.e.*,

$$\overline{I}_{\mathrm{dirty}}(u) = \langle \overline{V} \rangle (u). \quad (18)$$

Indeed, substituting $\langle \overline{V} \rangle (u)$ with the help of Eqs. 17 and 15 and taking the inverse Fourier transform, we get

$$\overline{I}_{\mathrm{dirty}}(u) = \overline{D}(u) \, \overline{I}(u), \quad (19)$$

$$\text{with} \quad \overline{D}(u) \equiv \int_{u_p} W(u_p, u - u_p) \, \overline{B}(u_p - u) \, \mathrm{d}u_p. \quad (20)$$

Here, I_{dirty} conforms to the usual idea of dirty image, *i.e.*, the convolution of a dirty beam by the sky brightness:

$$I_{\mathrm{dirty}}(\alpha) = \{D \star I\}(\alpha). \quad (21)$$

In contrast to the usual situation for single-field observations, the mix between a Fourier transform and a convolution of Eq. 1, associated with the specific processing[3] changes the image-plane measurement equation from a convolution of a dirty beam with the product $B \, I$ to a convolution of a dirty beam with I. *The dependency on the primary beam is still there. It is just transferred from a product of the sky brightness distribution into the definition of the dirty beam.*

3.1.2. Summary and interpretation

In summary, a theoretical implementation of wide-field synthesis implies

1. the possibility of Fourier transforming the visibility function along the α_s dimension (*i.e.* at constant u_p), which gives us a set of synthesized uv planes;
2. the possibility of shifting-and-averaging these synthesized uv planes to build the final, wide-field uv plane containing all the available information.

Using those tools, we are able to write the wide-field image-plane measurement equation as a convolution of a wide-field dirty beam (D) by the sky brightness (I), *i.e.*,

$$I_{\mathrm{dirty}}(\alpha) = \int_{\alpha'} D(\alpha - \alpha') \, I(\alpha') \, \mathrm{d}\alpha'. \quad (22)$$

We can write a convolution equation in this ideal case because the wide-field response of the instrument is shift-invariant; *i.e.*, D only depends on differences of the sky coordinates.

It is well-known that for a single-field observation, the dirty beam is the inverse Fourier transform of the sampling function. The shape of this sampling function is due to the combination of aperture synthesis (the interferometer antennas give a limited number of independent baselines) and Earth-rotation synthesis (the rotation of the Earth changes the projection of the physical baselines onto the plane perpendicular to the instantaneous line of sight). By analyzing via a Fourier transform, the evolution of the visibility function with the sky position, the Ekers & Rots scheme synthesizes visibilities at spatial frequencies needed to image a larger field of view than the interferometer primary beam. We thus propose to call this specific processing: *wide-field synthesis*.

3.2. General case: Incomplete sampling

Reality imposes limitations on the synthesis of spatial frequencies. Indeed, we have already stated that the visibility function is incompletely sampled both in the uv and sky planes. To take the sampling effects into account, we introduce the sampling function $S(u_p, \alpha_s)$, which is a sum of Dirac functions at measured positions[4]. The sampling function cannot be factored into the product of two functions, each only acting on one plane. Indeed, the Earth rotation happening during the source scanning implies a coupling of both dimensions of the sampling function. In other words, the uv coverage will vary with the scanned sky coordinate. This leads us to a shift-dependent situation, precluding us from writing the wide-field image-plane measurement equation as a true convolution. We nevertheless search for a wide-field image-plane measurement equation as close as possible to a convolution because all the inversion methods devised in the past three decades in radioastronomy are tuned to deconvolve images. The simplest mathematical way to generalize Eq. 22 to a shift-dependent situation is to write it as

$$I_{\mathrm{dirty}}(\alpha) = \int_{\alpha'} D(\alpha - \alpha', \alpha) \, I(\alpha') \, \mathrm{d}\alpha'. \quad (23)$$

---

[3] *i.e.* direct Fourier transform along the α_s dimension, shift-and-average to define a final wide-field uv plane, and inverse Fourier transform.

[4] Loosely speaking, the sampling function can be thought as a function whose value is 1 where there is a measure and 0 elsewhere.

In this section, we show how the *linear character of the imaging process* allows us to do this. Section 3.2.1 derives the impact of incomplete sampling on the Ekers & Rots equation, and Sect. 3.2.2 derives the wide-field measurement equation in the uv plane. Section 3.2.3 interprets these results.

3.2.1. Effect on the Ekers & Rots equation

The sampled visibility function, SV, is defined as the product of S and V and $\overline{SV}$ its Fourier transform along α_s, *i.e.*,

$$SV(u_p, \alpha_s) \equiv S(u_p, \alpha_s) \, V(u_p, \alpha_s), \tag{24}$$

$$\text{and} \quad SV(u_p, \alpha_s) \overset{\alpha_s}{\underset{u_s}{\supset}} \overline{SV}(u_p, u_s). \tag{25}$$

Because SV_{u_p} is the product of two functions of α_s, we can use the convolution theorem to show that $\overline{SV}_{u_p}$ is the convolution of $\overline{S}_{u_p}$ by $\overline{V}_{u_p}$, *i.e.*,

$$\overline{SV}_{u_p}(u_s) = \int_{u_s'} \overline{S}_{u_p}(u_s - u_s') \, \overline{V}_{u_p}(u_s') \, du_s'. \tag{26}$$

By replacing $\overline{V}_{u_p}$ with the help of the Ekers & Rots relation (Eq. 15), we derive

$$\overline{SV}_{u_p}(u_s) = \int_{u_s'} \overline{S}_{u_p}(u_s - u_s') \, \overline{B}(-u_s') \, \overline{I}(u_p + u_s') \, du_s'. \tag{27}$$

As $\overline{B}$ is bounded inside the $[-d_{\text{prim}}, +d_{\text{prim}}]$ interval, $\overline{SV}_{u_p}(u_s)$ is a *local* average, weighted by $\overline{S}_{u_p}(u_s - u_s') \, \overline{B}(-u_s')$, of $\overline{I}(u_p + u_s')$ around the u_p spatial frequency.
 As expected, we recover Eq. 15 for the ideal case (*i.e.*, infinite, continuous visibility function) because then $\overline{S}_{u_p}(u_s - u_s') = \delta(u_s - u_s')$. A more interesting case arises when the visibility function is continuously sampled over a limited sky field of view, *i.e.*,

$$\forall u_p, \qquad S(u_p, \alpha_s) = 1 \quad \text{if} \quad |\alpha_s| \leq \theta_{\text{field}}/2, \tag{28}$$

$$\forall u_p, \qquad S(u_p, \alpha_s) = 0 \quad \text{if} \quad |\alpha_s| > \theta_{\text{field}}/2. \tag{29}$$

After Fourier transform this gives

$$\forall u_p, \quad \overline{S}(u_p, u_s) = \frac{1}{d_{\text{field}}} \operatorname{sinc}\left(\frac{u_s}{d_{\text{field}}}\right). \tag{30}$$

In this case, the local average of the sky brightness Fourier components happens on a typical uv scale equal to d_{field}. However, the sinc function is known to decay only slowly. Some observing strategy (e.g. quickly observing outside the edges of the targeted field of view to provide a bandguard) could be considered to apodize the sky-plane dependence of the sampling function, resulting in faster decaying $\overline{S}$ functions, hence in less mixing of the wide-field spatial frequencies.

3.2.2. uv-plane wide-field measurement equation

Because we aim at estimating the Fourier component of $\overline{I}$, we introduce the following change of variables $u' \equiv u_p + u_s'$ and $du' = du_s'$, to derive

$$\overline{SV}_{u_p}(u_s) = \int_{u'} \overline{S}_{u_p}(u_p + u_s - u') \, \overline{B}(u_p - u') \, \overline{I}(u') \, du'. \tag{31}$$

We then shift-and-average $\overline{SV}(u_p, u_s)$ to build the Fourier transform of a wide-field dirty image

$$\overline{I}_{\text{dirty}}(u) \equiv \langle \overline{SV} \rangle(u), \quad \text{with} \quad u = u_p + u_s. \tag{32}$$

Substituting the shift-and-average operator by its definition and using Eq. 31 to replace $\overline{SV}_{u_p}(u_s)$, we derive

$$\overline{I}_{\text{dirty}}(u)$$

$$= \iint_{u_p u'} W(u_p, u - u_p) \, \overline{S}(u_p, u - u') \, \overline{B}(u_p - u') \, \overline{I}(u') \, du_p du'. \tag{33}$$

This uv-plane wide-field measurement equation can be written as

$$\overline{I}_{\text{dirty}}(u) = \int_{u'} \overline{D}(u', u - u') \, \overline{I}(u') \, du', \tag{34}$$

if we enforce the following equality

$$\overline{D}(u', u - u') \equiv \int_{u_p} W(u_p, u - u_p) \, \overline{S}(u_p, u - u') \, \overline{B}(u_p - u') \, du_p. \tag{35}$$

This is one way to define $\overline{D}$, which is convenient though unusual. It is implicit in this definition that we need to make a change of variable ($u'' = u - u'$) to derive

$$\overline{D}(u', u'') \equiv \int_{u_p} W(u_p, u' + u'' - u_p) \, \overline{S}(u_p, u'') \, \overline{B}(u_p - u') \, du_p. \tag{36}$$

In the following, we use either one or the other definition of $\overline{D}$, depending on convenience.

3.2.3. Interpretation

Appendix A.2 demonstrates that the image and uv-plane wide-field measurement equations (Eqs. 23 and 34) are equivalent if

$$D(\alpha_p, \alpha_s) \overset{(\alpha_p, \alpha_s)}{\underset{(u_p, u_s)}{\supset}} \overline{D}(u_p, u_s). \tag{37}$$

The image-plane wide-field measurement equation (Eq. 23) can be written as

$$I_{\text{dirty}}(\alpha) = \{D^\alpha \star I\}(\alpha). \tag{38}$$

Its interpretation is straightforward: The sky brightness distribution is convolved with a dirty beam, $D(\alpha', \alpha'')$, which varies with the sky coordinate α''. This raises the question of the rate of change of the dirty beam with the sky coordinate. This question is addressed in Sects. 4.2 and 5.

4. Gridding by convolution and regular resampling

We want to Fourier transform the raw visibilities along the sky dimension (α_s) at some constant value in the u_p dimension. The raw data, however, is sampled on an irregular grid in both the uv and sky planes. We need to grid the measured visibilities in both the uv and the sky planes before Fourier transformation for different reasons. First, the gridding in the uv plane will handle the variation in the spatial frequency as the sky is scanned, *i.e.*, the difficulty and perhaps the impossibility of Fourier-transforming at a completely constant u_p value. Second, the gridding along the sky dimension allows the use of Fast Fourier Transforms. As usual, we grid through convolution and regular resampling.

4.1. Convolution

4.1.1. Definitions

We first define a gridding kernel that depends on both dimensions, $\mathcal{G}(u, \alpha_s)$. This gridding kernel can be chosen as the product of two functions, simplifying the following demonstrations:

$$\mathcal{G}(u_p, \alpha_s) \equiv g(u_p)\, \gamma(\alpha_s). \tag{39}$$

We then define the sampled visibility function gridded in both the uv and sky planes as

$$SV^{\mathcal{G}}(u_p, \alpha_s) \equiv \{\mathcal{G} \star SV\}(u_p, \alpha_s) \tag{40}$$

$$= \iint_{u'_p \alpha'_s} g(u_p - u'_p)\, \gamma(\alpha_s - \alpha'_s)\, SV(u'_p, \alpha'_s)\, \mathrm{d}u'_p \mathrm{d}\alpha'_s. \tag{41}$$

Finally, when assessing the impact of the gridding on the measurement equation 34, a new function,

$$\Sigma(u_p, \alpha_s, \alpha''_s) \equiv S(u_p, \alpha_s)\, B(\alpha''_s - \alpha_s), \tag{42}$$

and its Fourier transforms naturally appear in the equations. Defining the following Fourier transform relationships

$$\Sigma(u_p, \alpha_s, \alpha''_s) \overset{\alpha''_s}{\underset{u''_s}{\supset}} \overline{\Sigma}(u_p, u_s, \alpha''_s), \tag{43}$$

and

$$\overline{\Sigma}(u_p, u_s, \alpha''_s) \overset{\alpha_s}{\underset{u_s}{\supset}} \underline{\overline{\Sigma}}(u_p, u_s, u''_s), \tag{44}$$

we easily derive

$$\overline{\Sigma}(u_p, \alpha_s, u''_s) = S(u_p, \alpha_s)\, \overline{B}(u''_s)\, \mathrm{e}^{-i2\pi u''_s \alpha_s}, \tag{45}$$

and

$$\underline{\overline{\Sigma}}(u_p, u_s, u''_s) = \overline{S}(u_p, u_s + u''_s)\, \overline{B}(u''_s). \tag{46}$$

Using these notations, we have before gridding,

$$SV(u_p, \alpha_s) = \int_{\alpha_p} \Sigma(u_p, \alpha_s, \alpha_p)\, I(\alpha_p)\, \mathrm{e}^{-i2\pi\alpha_p u_p}\, \mathrm{d}\alpha_p, \tag{47}$$

and

$$\underline{\overline{D}}(u', u - u') = \int_{u_p} W(u_p, u - u_p)\, \underline{\overline{\Sigma}}(u_p, u - u_p, u_p - u')\, \mathrm{d}u_p. \tag{48}$$

4.1.2. Conservation of the wide-field measurement equation

Appendix A.3 demonstrates that the wide-field dirty image is here again the convolution of the sky brightness I by a wide-field dirty beam D^α or, in the Fourier plane,

$$\overline{I}^{\mathcal{G}}_{\mathrm{dirty}}(u) \equiv \left\langle \overline{SV^{\mathcal{G}}} \right\rangle(u) = \int_{u'} \underline{\overline{D}}^{\mathcal{G}}(u', u - u')\, \overline{I}(u')\, \mathrm{d}u' \tag{49}$$

with

$$\underline{\overline{D}}^{\mathcal{G}}(u', u - u') \equiv \int_{u_p} W(u_p, u - u_p)\, \underline{\overline{\Sigma}}^{\mathcal{G}}(u_p, u - u_p, u')\, \mathrm{d}u_p, \tag{50}$$

where $\overline{\Sigma}^{\mathcal{G}}(u_p, \alpha_s, u')$

$$\equiv \iint_{u'_p \alpha'_s} g(u_p - u'_p)\, \gamma(\alpha_s - \alpha'_s)\, \overline{\Sigma}(u'_p, \alpha'_s, u'_p - u')\, \mathrm{d}u'_p \mathrm{d}\alpha'_s. \tag{51}$$

We thus have equations that resemble those containing the sampling function alone, except for 1) the replacement of the generalized sampling function $\overline{\Sigma}$ by its gridded version $\overline{\Sigma}^{\mathcal{G}}$ and 2) the way the variables are linked together both in the gridding of $\overline{\Sigma}$ (i.e., Eq. 51) and in the averaging of $\underline{\overline{\Sigma}}^{\mathcal{G}}$ (i.e., Eq. 50).

4.2. Regular resampling

It is well known that too low a resampling rate in one space implies power aliasing in the conjugate space (See e.g. Bracewell 2000; Press et al. 1992). Aliasing must be avoided as much as possible because it folds power outside the imaged region back into it. Table 3 defines the intervals of definition of the different functions we are dealing with (i.e., visibilities, primary beam, dirty image, and dirty beam), as well as the associated sampling rates needed to enforce Nyquist sampling. The boundary values of the definition intervals ($|u|_{\mathrm{max}}$ and $|\alpha|_{\mathrm{max}}$) are related to the sampling rates ($\partial\alpha$ and ∂u, respectively) through

$$|u|_{\mathrm{max}}.\partial\alpha = |\alpha|_{\mathrm{max}}.\partial u = \frac{1}{n_{\mathrm{samp}}}, \tag{52}$$

where n_{samp} is an integer characterizing the sampling. Nyquist sampling implies $n_{\mathrm{samp}} = 2$. However, slight oversampling (e.g. $n_{\mathrm{samp}} = 3$) is often recommended because the measures suffer from errors and the deconvolution is a nonlinear process. In this section, we examine the properties of the different functions to define their associated sampling rates.

4.2.1. The α_s sampling rate of the visibility function

When Fourier transforming the measurement Eq. 1 along the α_s axis, we derive the Ekers & Rots equation (15). This equation implies that $\overline{V}(u_p, u_s)$ is bounded inside the $[-d_{\mathrm{prim}}, +d_{\mathrm{prim}}]$ spatial frequency interval along the u_s axis. As a result, the visibility function needs to be regularly resampled at a rate of only $0.5/d_{\mathrm{prim}}$ to satisfy the Nyquist theorem. This was first pointed out by Cornwell (1988). This sampling rate is equal to $\theta_{\mathrm{prim}}/2$ or $\sim \theta_{\mathrm{fwhm}}/2.4$. The "usual, wrong" habit of sampling at $\theta_{\mathrm{fwhm}}/2$ is indeed undersampling with aliasing as a consequence. Mangum et al. (2007) discuss the consequences of undersampling in-depth in the framework of single-dish imaging.

4.2.2. The u_p sampling rate of the visibility function

Now, the Fourier transform of the measurement equation 1 along the u_p axis gives

$$\underset{\sim}{V}(\alpha_p, \alpha_s) = B(\alpha_p - \alpha_s)\, I(\alpha_p), \tag{53}$$

where $\underset{\sim}{V}(\alpha_p, \alpha_s) \overset{\alpha_p}{\underset{u_p}{\supset}} V(u_p, \alpha_s). \tag{54}$

We use the tilde sign under V to denote the inverse Fourier transform of V along its first dimension. A well-known Fourier transform property implies that B has infinite support because $\overline{B}$ is bounded. The resampling rate along the u_p axis therefore depends on the properties of the product of $B(\alpha_p - \alpha_s)$ times $I(\alpha_p)$ as a function of α_p. While no unique answer exists, three facts help us to find the right

Table 3. Interval ranges of definition and associated sampling rates for the used functions.

Functions	Intervals	Samplings
Visibilities	$|u_p| \leq d_{\max}$	$\partial u_p = 2\, d_{\mathrm{alias}}/n_{\mathrm{samp}}$
	$|\alpha_p| \leq \theta_{\mathrm{alias}}/2$	$\partial \alpha_p = \theta_{\mathrm{syn}}/n_{\mathrm{samp}}$
	$|u_s| \leq d_{\mathrm{prim}}$	$\partial u_s = 2\, d_{\mathrm{image}}/n_{\mathrm{samp}}$
	$|\alpha_s| \leq \theta_{\mathrm{image}}/2$	$\partial \alpha_s = \theta_{\mathrm{prim}}/n_{\mathrm{samp}}$
Primary beam	$|u'_s| \leq d_{\mathrm{prim}}$	$\partial u'_s = 2\, d_{\mathrm{alias}}/n_{\mathrm{samp}}$
	$|\alpha'_s| \leq \theta_{\mathrm{alias}}/2$	$\partial \alpha'_s = \theta_{\mathrm{prim}}/n_{\mathrm{samp}}$
	$|u''_s| \leq d_{\mathrm{prim}}$	$\partial u''_s = 2\, d_{\mathrm{alias}}/n_{\mathrm{samp}}$
	$|\alpha''_s| \leq \theta_{\mathrm{alias}}/2$	$\partial \alpha''_s = \theta_{\mathrm{prim}}/n_{\mathrm{samp}}$
Dirty image	$|u| \leq d_{\max}$	$\partial u = 2\, d_{\mathrm{image}}/n_{\mathrm{samp}}$
	$|\alpha| \leq \theta_{\mathrm{image}}/2$	$\partial \alpha = \theta_{\mathrm{syn}}/n_{\mathrm{samp}}$
Dirty beam	$|u'| \leq d_{\max}$	$\partial u' = 2\, d_{\mathrm{image}}/n_{\mathrm{samp}}$
	$|\alpha'| \leq \theta_{\mathrm{image}}/2$	$\partial \alpha' = \theta_{\mathrm{syn}}/n_{\mathrm{samp}}$
	$|u''| \leq d_{\mathrm{prim}}$	$\partial u'' = 2\, d_{\mathrm{image}}/n_{\mathrm{samp}}$
	$|\alpha''| \leq \theta_{\mathrm{image}}/2$	$\partial \alpha'' = \theta_{\mathrm{prim}}/n_{\mathrm{samp}}$

Table 4. Minimum sizes of the dirty beam images to get an image fidelity or a dynamic range greater than a given value.

Minimum fidelity	$\theta_{\mathrm{alias}}/\theta_{\mathrm{fwhm}}$[a]		
or dynamic range	$(f_{\mathrm{b}} = 0)$[b]	$(f_{\mathrm{b}} = 0.0625)$	$(f_{\mathrm{b}} = 0.1)$
10^2	2.2	2.2	2.2
10^3	3.5	3.7	6.6
10^4	8.4	13.4	13.7
10^5	19.8	> 20.0	> 20.0

Notes. [a] The image sizes are expressed in units of the primary beam full width at half maxium. [b] The computation is done for 3 different ratios of the secondary-to-primary diameters (*i.e.* f_{b}, the antenna blockage factors). The values are derived from the modeling of the antenna power patterns shown in Fig. 3.

sampling rate: 1) B falls off relatively quickly. 2) The result depends on the spatial distribution of the sky brightness and in particular on the dynamic range in brightness needed to accurately image it. 3) The measure of $\underset{\sim}{V}(\alpha_p, \alpha_s)$ has a limited accuracy owing to thermal noise, phase noise, and other possible systematics (e.g. pointing errors). For simplicity, we quantify the measurement accuracy by a single number, namely the maximum instrumental fidelity measured in the image plane as defined in Pety et al. (2001). There are two cases:

1. The maximum instrumental fidelity limits the dynamic range in brightness. For instance, Pety et al. (2001) showed that the fidelity of interferometric imaging at (sub)-millimeter wavelengths will be limited to a few hundred. In this case, $\underset{\sim}{V}(\alpha_p, \alpha_s)$ aliasing can be tolerated when the amplitude of B is less than a fraction of the inverse of the maximum instrumental fidelity.
2. The maximum instrumental fidelity is much greater than the image fidelity, as can be the case at centimeter wavelengths. In this case, $\underset{\sim}{V}(\alpha_p, \alpha_s)$ aliasing can only be tolerated when the amplitude of B is less than a fraction of the inverse of the dynamic range of the image.

The criterion derived in each case gives a typical image size (θ_{alias}), which can be converted into the desired u_p sampling rate. To be more quantitative, Fig. 3 models the
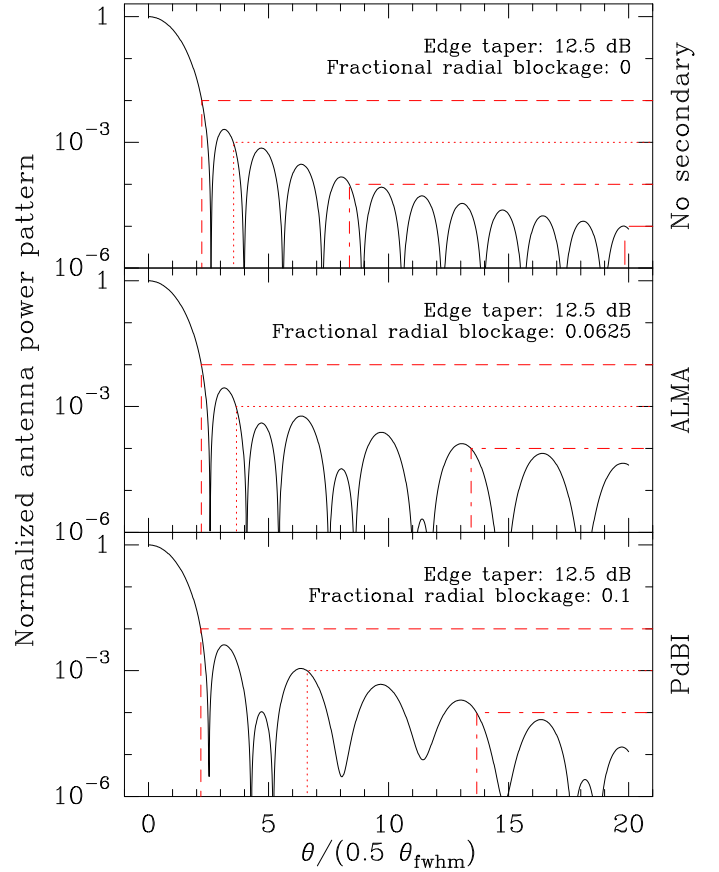
Fig. 3. Simple models of the antenna power patterns as a function of the sky angle in units of half the primary beam FWHM (θ_{fwhm}). In the 3 cases shown, the illumination is Gaussian with an edge taper of 12.5 dB but 3 different ratios of the secondary-to-primary diameters (*i.e.* f_{b}, the antenna blockage factors) are considered (see e.g. Goldsmith 1998, chapter 6). The middle and bottom panels respectively model ALMA and PdBI antennas. The red lines define the minimum angular sizes for which the antenna power pattern is less than a given fraction.

normalized antenna power patterns of an antenna illuminated by a Gaussian beam of 12.5 dB edge taper and with a given blockage factor (ratio of the secondary-to-primary diameters). The top panel presents an ideal case without secondary miror, while the middle and bottom panels present simple models of the ALMA and PdBI antennas. The largest angular sizes at which the power patterns are less than a given value, $\mathcal{P}_0$, is a first-order estimate of $\theta_{\mathrm{alias}}/2$ to get a fidelity or dynamic range higher than $1/\mathcal{P}_0$. Table 4 gives the values of $\theta_{\mathrm{alias}}/\theta_{\mathrm{fwhm}}$ as a function of the searched fidelity or dynamic range. This condition is sufficient but not necessary. Indeed, the aliasing properties also depend on the brightness distribution of the source.

4.2.3. The u sampling rate of $\overline{I}_{\mathrm{dirty}}(u)$

We have no garantee that the sky outside the targeted field of view is devoid of signal, so the only way to ensure a given dynamic range inside the targeted field of view is to choose the image size large enough so that the aliasing of potential outside sources is negligible. This means that the

dirty image size must be equal to the field-of-view size plus the tolerable aliasing size

$$\theta_{\text{image}} = \theta_{\text{field}} + \theta_{\text{alias}}. \tag{55}$$

The conjugate uv distance and associated uv sampling then are

$$d_{\text{image}} = \frac{d_{\text{field}}}{1 + \frac{d_{\text{field}}}{d_{\text{alias}}}} \quad \text{and} \quad \partial u = \frac{d_{\text{image}}}{n_{\text{samp}}}. \tag{56}$$

4.2.4. The u' and u'' sampling rates of $\overline{D}(u',u'')$

The u'' axis must thus be sampled at the same rate as the second dimension of the definition space of $\overline{S}$, i.e., as u_s. Moreover, u' has in this equation a behavior ($u' = u_p + u''_s$) similar to u ($= u_p + u_s$). It must thus have the same sampling behavior as u. This sampling rate ($\partial u' = d_{\text{image}}/n_{\text{samp}}$) is quite high. Some deconvolution methods (see below) allow us to relax this sampling rate.

4.3. Absence of gridding "correction"

Imaging of single-field observations goes through the following steps: 1) convolution by a gridding kernel, 2) regular resampling, 3) fast Fourier transform, and 4) gridding "correction". The so-called gridding "correction" is a division of the dirty beam and dirty image by the Fourier transform of the gridding kernel used in the initial convolution. This step is mandatory when imaging single-field observations to keep the image-plane measurement equation as a simple convolution equation (see e.g. Sramek & Schwab 1989). When imaging wide-field observations, as proposed here, the Fourier transform along the α_s dimension, followed by the shift-and-average operation, freeze the convolution kernel into the dirty beam of the wide-field measurement equation. This is why the gridding "correction" step is irrelevant here.

5. Dirty beams, weighting, and deconvolution

In radioastronomy, the dirty beam is the response of the interferometer to a point source. In the wide-field synthesis framework, the response of the interferometer to a point source, D, a priori depends on the source position on the sky. $D(\alpha',\alpha'')$ can thus be interpreted as a set of dirty beams, with each dirty beam referred to by its fixed α'' sky coordinate. These simple facts raise several questions. What are the properties of the convolution kernel? Is it possible to modify these properties? How do we deconvolve the dirty image?

5.1. A set of wide-field dirty beams

With the wide-field synthesis framework proposed here, Appendix A.6 shows that
$D(\alpha', \alpha'') =$

$$\iint_{\alpha_p \alpha_s} B(\alpha'' - \alpha' - \alpha_s)\Omega(\alpha' - \alpha_p, \alpha'' - \alpha_s)\Delta(\alpha_p, \alpha_s)\, d\alpha_p d\alpha_s. \tag{57}$$

$$\text{where} \quad \Delta(\alpha_p, \alpha_s) \overset{\alpha_p}{\underset{u_p}{\supset}} S(u_p, \alpha_s), \tag{58}$$

$$\text{and} \quad \Omega(\alpha', \alpha'') \overset{(\alpha',\alpha'')}{\underset{(u',u'')}{\supset}} W(u', u''). \tag{59}$$

$\Delta(\alpha_p,\alpha_s)$ is the single-field dirty beam, associated with the uv sampling at the sky coordinate α_s. And $\Omega(\alpha',\alpha'')$ will be called the image plane weighting function, while $W(u',u'')$ is the uv plane weighting function. The set of wide-field dirty beams D is then the double convolution of the image plane weighting function and the single-field dirty beams, apodized by the primary beam at the current sky position α_s.

While the shape of the single-field dirty beam is directly given by the Fourier transform of the sampling function, the shape of the wide-field dirty beam depends, directly or through Fourier transforms, on the sampling function (S), the primary beam shape (B), and the weighting function (W). Moreover, the wide-field dirty beam shape a priori varies slowly with the sky position, since it is basically constant over the primary beamwidth as stated in Sect. 4.2. It nevertheless varies, implying, for instance, a "slow" variation of the synthesized resolution over the whole field of view.

While the single-field and wide-field dirty beam expressions seem very different, they share the same property of expressing the way the interferometer is used to synthesize a telescope of larger diameter in the image plane. In other words, the sampling function for single-field imaging and $\overline{D}$ for wide-field imaging express the sensitivity of the interferometer to a given spatial frequency. These uv space functions are called the transfer functions of the interferometer (Thompson et al. 1986, chapter 5). Modifying the transfer function has a direct impact on the measured quantity. Once the interferometer is designed and the observations are done, the only way to change this transfer function is data weighting.

An ideal set of wide-field dirty beams, $D(\alpha',\alpha'')$, would have the following properties. All the wide-field dirty beams should be identical (i.e., independent of the α'' sky coordinate) and equal to a narrow Gaussian (its FWHM giving the image resolution). *This would give the product of a wide Gaussian of u' by a Dirac function of u'', as the ideal wide-field transfer function, $\overline{D}(u',u'')$.*

5.2. Dirty beam shapes and weighting

When imaging single-field observations, giving a multiplicative weight to each visibility sample is an easy way to modify the shape of the dirty beam and thus the properties of the dirty and deconvolved images. Natural weighting (which maximizes signal-to-noise ratio), robust weighting (which maximizes resolution), and tapering (which enhances brightness sensitivity at the cost of a lower resolution) are the most popular weighting techniques (see e.g. Sramek & Schwab 1989).

In the case of wide-field synthesis, a multiplicative weight can also be attributed to each visibility sample before any processing. However, the weighting is also at the heart of the wide-field synthesis because it is an essential part of the shift-and-average operation. No constraint has been set on the weighting function up to this point, which indicates that the weighting function (W) gives us a degree of freedom in the imaging process. We look in turn at both kinds of weighting. In both cases, an obvious issue is the definition of the optimum weighting functions. As in the

case of single-field imaging, there is no single answer to this question. It depends on the conditions of the observation and on the imaging goals.

5.2.1. Weighting the measured visibilities

Natural weighting consists of slightly changing the definition of the sampling function. It is now set to a normalized natural weight where there is a measure and 0 elsewhere. The natural weight is usually defined as the inverse of the thermal noise variance, computed from the radiometric equation, *i.e.*, from the system temperature, the frequency resolution, and the integration time. Using this weighting scheme before computing the first Fourier transform along the α_s sky dimension makes sense because the observing conditions (and thus the noise) vary from visibility to visibility.

We propose to generalize this weighting scheme to other observing conditions than just the system noise. Indeed, critical limitations of interferometric wide-field imaging are pointing errors, tracking errors, atmospheric phase noise (in the (sub)-millimeter domain), etc. While techniques exist for coping with these problems (e.g., water vapor radiometer, direction-dependent gains: Bhatnagar et al. 2008), they are not perfect. The usual way to deal with the remaining problems is to flag the source data based on *a priori* knowledge of the problems, e.g., pointing measurement, tracking errors, rms phase noise on calibrators, etc. However, flagging involves the definition of thresholds, while reality is never black and white. It can thus be asked whether some weighting scheme could be devised to minimize the effect of pointing errors, tracking errors or phase noise on the resulting image. We propose to modulate natural weighting based on the *a priori* knowledge of the observing conditions.

5.2.2. Weighting the synthesized visibilities

Robust weighting or tapering the *measured* visibilities do not make sense in wide-field synthesis because the dirty image is made from the synthesized visibilities after the first Fourier transform along the α_s sky dimension. A weighting function W then appears naturally as part of the shift-and-average operator. Its optimum value depends on the properties of the measured sampling function. Here are a few examples.

Infinite, continuous sampling. This is the ideal case studied in Sect. 3.1. Knowing that the Ekers & Rots equation (15) links the quantity we want to estimate, *i.e.*, $\overline{I}$, to many noisy[5] measurements, $\overline{V}(u_p, u_s)$, via a product by $\overline{B}$ (assumed to be perfectly defined), we can invoke a simple least-squares argument (see e.g. Bevington & Robinson 2003) to demonstrate that the optimum weighting function is

$$W(u_p, u - u_p) = \frac{w(u_p, u - u_p)\,\overline{B}(u_p - u)}{\int_{u_p} w(u_p, u - u_p)\,\overline{B}^2(u_p - u)\,\mathrm{d}u_p}, \quad (60)$$

with $w(u_p, u_s)$ the weight computed from the inverse of the noise variance of $\overline{V}(u_p, u_s)$. Using Eq. 20, it is then easy to demonstrate that $\overline{D}(u) = 1$, and then

---

[5] The noise is assumed to have a Gaussian probability distribution function.

$I_{\mathrm{dirty}}(\alpha) = I(\alpha)$. The dirty image is a direct estimate of the sky brightness; *i.e.*, deconvolution is superfluous.

Complete sampling. The signal is Nyquist-sampled, but it has a finite support in both the uv and sky planes, implying a finite synthesized resolution and a finite field of view. In contrast to the previous case, this one may have practical applications, e.g., observations done with ALMA in its compact configuration. Indeed, the large number of independent baselines coupled to the design of the ALMA compact configuration ensure a complete, almost Gaussian, sampling for each snapshot. In this case, the best choice may be to choose the weighting function so that all the dirty beams are identical to the same Gaussian function. In this case, the deconvolution would also be superfluous.

Incomplete sampling. This is the more general case studied in Sect. 3.2. The signal not only has a finite support but it also is undersampled (at least in the uv plane). The deconvolution is mandatory. The choice of the weighting function thus will depend on imaging goals.

If the user needs the best signal-to-noise ratio, some kind of natural weighting will be needed. It is tempting to use Eq. 60 as a natural weighting scheme. However, the main condition for derivation of this weighting function, *i.e.*, the Ekers & Rots equation (15), is not valid anymore, as the noisy measured quantity $(\overline{SV})$ is now linked to the quantity we want to estimate $(\overline{I})$ by a local average (see Eq. 31). This is why it was more appropriate to try to get a Gaussian dirty beam shape in the complete sampling case.

If the signal-to-noise ratio is high enough, the user has two choices. Either he/she wants to maximize angular resolution power and needs some kind of robust weighting, or he/she wants to get the more homogeneous dirty beam shape over the whole field of view. This requirement cannot always be fully met. The Ekers & Rots scheme enables us to recover unmeasured spatial frequencies only in regions near to measured ones, because $\overline{B}$ has a finite support.

5.3. Deconvolution

Writing the image-plane measurement equation in a convolution-like way is very interesting because all the deconvolution methods developed in the past 30 years are optimized to treat deconvolution problems (see e.g. Högbom 1974; Clark 1980; Schwab 1984; Narayan & Nityananda 1986). For instance, it should be possible to deconvolve Eq. 23 with just slight modifications to the standard CLEAN algorithms. Indeed, Eq. 23 can be interpreted as the convolution of the sky brightness by a set of dirty beams, so that the only change, once a CLEAN component is found, would be the need to find the right dirty beam in this set in order to remove the CLEAN component from the residual image.

Following Clark (1980) and Schwab (1984), most algorithms today deconvolve in alternate minor and major cycles. During a minor cycle, a solution of the deconvolution is sought with a simplified (hence approximate) dirty beam. During a major cycle, the current solution is subtracted either from the original dirty image using the exact dirty beam or from the measured visibilities, implying a new gridding step. In both cases, the major cycles result in greater accuracy. The iteration of minor and major cycles

enables one to find an accurate solution with better computing efficiency. In our case, the approximate dirty beams used in the minor cycle could be 1) dirty beams of a much smaller size than the image, or 2) a reduced set of dirty beams (*i.e.*, guessing that the typical variation sizescale of the dirty beams with the sky coordinate is much larger than the primary beamwidth), or 3) both simultaneously. The model would be subtracted from the original visibilities before re-imaging at each major cycle. The trade-off is between the memory space needed to store a full set of accurate dirty beams and the time needed to image the data at each major step. Some quantitative analysis is needed to know how far the dirty beams can be approximated in the minor cycle.

It is worth noting that the accuracy of the deconvolved image will be affected by edge effects. Indeed, the dirty brightness at the edges of the observed field of view is attenuated by the primary beam shape. When deconvolving these edges, the deconvolved brightness will be less precise, because the primary beam has a low amplitude there. This only affects the edges, because inside the field of view, every sky position should be observed a fraction of the time with a primary beam amplitude between 0.5 and 1. This edge effect is nevertheless expected to be much less troublesome than the inhomogeneous noise level resulting from standard mosaicking imaging (see Sect. 7.1).

6. Short spacings

6.1. The missing flux problem

Radio interferometers are bandpass instruments; *i.e.*, they filter out not only the spacings longer than the largest baseline length but also the spacings shorter than the shortest baseline length, which is typically comparable to the diameter of the interferometer antennas. In particular, radio interferometers do *not* measure the visibility at the center of the uv plane (the so-called "zero spacing"), which is the total flux of the source in the measured field of view.

The lack of short baselines or *short spacings* has strong effects as soon as the size of the source is more than about 1/3 to 1/2 of the interferometer primary beam. Indeed, when the size of the source is small compared to the primary beam of the interferometer, the deconvolution algorithms use, in one way or another, the information of the flux at the lowest *measured* spatial frequencies for extrapolating the total flux of the source. The extreme case is a point source at the phase center for which the amplitude of all the visibilities is constant and equal to the total flux of the source: Extrapolation is then exact. However, the larger the size of the source, the worse the extrapolation, which then underestimates the total source flux. This is the well-known problem of the missing flux that observers sometimes note when comparing a source flux measured by a mm interferometer with the flux observed with a single-dish antenna.

Wide-field synthesis does *not* recover the full short spacings. Let us assume that the visibility function is continuously sampled from d_{min} to d_{max}, with $d_{min} \sim 1.5\,d_{prim}$. The length of the averaging linepath[6], $L(u)$, can be interpreted as the number of measures that contribute to the estimation of $\overline{I}(u)$. Figure 4 shows the variations of $L(u)$

---

[6] The notion of averaging linepath has been introduced in Sect. 3.1 (see in particular Eq. 16).
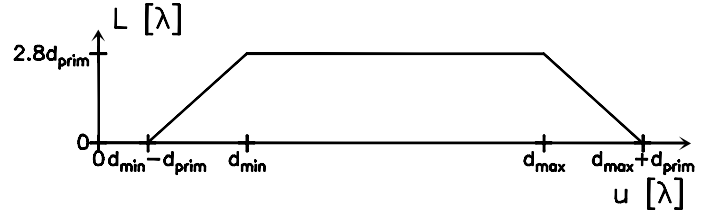
Fig. 4. Length of the averaging linepaths displayed as black lines in panel c.2) of Fig. 2, as a function of the spatial scale in the final, wide-field uv plane. In the case of a continuous sampling of u_p between d_{min} and d_{max}, these quantities can be interpreted as the number of measures that contribute to the estimate of $\overline{I}(u)$.

function when starting from a visibility function continuously defined in the $[d_{min}, d_{max}]$ interval along the u_p dimension. We can expect to recover $\overline{I}(u)$ only inside the $[d_{min} - d_{prim}, d_{max} + d_{prim}]$ interval. In particular, information on short spacings lower than $d_{min} - d_{prim}$ (e.g. the crucial zero spacing) cannot be recovered when using a homegeneous interferometer, and the short spacings in the interval $[d_{min} - d_{prim}, d_{min}]$ are recovered with increasing accuracy from $d_{min} - d_{prim}$ to d_{min}. Both effects imply the need for complementary instruments to accurately measure the missing short-spacings.

6.2. Usual hardware and software solutions

To derive the correct result for larger source sizes, it is necessary to complement the interferometer data with additional data, which contain the missing short-spacing information. The IRAM-30m single-dish telescope is used to complement the Plateau de Bure Interferometer. Short-spacing information can also be in part recovered with a secondary array of smaller antennas and shorter baselines (e.g. the CARMA interferometer). In the ALMA project, the short-spacing information will be derived by a combination of four 12m-single-dish antennas and an interferometer of 12 antennas of 7 meters called ACA (Atacama Compact Array).

From the software point-of-view, two main families of algorithms exist in the standard processing of mosaics. Either the short-spacing information is combined on the deconvolved image (*i.e.*, the interferometer data is imaged and deconvolved separately) through a *hybridization* in the Fourier plane (see e.g. Pety et al. 2001), or the long and short-spacing information is imaged and/or deconvolved jointly. In this category, we find the *pseudo-visibility* technique, which produces interferometric-like visibilities from single-dish maps (see e.g. Pety et al. 2001; Rodríguez-Fernández et al. 2008, and references therein), and the multi-resolution deconvolution algorithms, which work on images containing different spatial frequency ranges.

In the next two sections, we show how wide-field synthesis naturally processes the short-spacing information either from single-dish or from heterogeneous arrays.

6.3. Processing short spacings from single-dish measurements

The single-dish measurement equation can be written as

$$I_{\rm sd}(\alpha) = S_{\rm sd}(\alpha) \int_{\alpha'} B_{\rm sd}(\alpha' - \alpha) \, I(\alpha') \, {\rm d}\alpha', \qquad (61)$$

where $I_{\rm sd}$ is the measured single-dish intensity, $S_{\rm sd}$ the single-dish sampling function, and $B_{\rm sd}$ the single-dish antenna power pattern. As already stated in the introduction, the above integral is identical to the ideal measurement equation of interferometric wide-field imaging taken in $u_p = 0$. If we define a single-dish visibility function as

$$V_{\rm sd}(u_p = 0, \alpha) \equiv \int_{\alpha'} B_{\rm sd}(\alpha' - \alpha) \, I(\alpha') \, {\rm d}\alpha', \qquad (62)$$

we can thus write the measured single-dish intensity as

$$I_{\rm sd}(\alpha) = S_{\rm sd}(\alpha) \, V_{\rm sd}(u_p = 0, \alpha). \qquad (63)$$

The recognition that the single-dish measurement equation is a particular case of the interferometric wide-field measurement equation opens the way to treating both the single-dish and interferometric data sets through exactly the same processing steps. We just have to define a hybrid sampling function, $S_{\rm hyb}$, as

$$S_{\rm hyb}(u_p \neq 0, \alpha) = S(u_p, \alpha) \qquad (64)$$
$$S_{\rm hyb}(u_p = 0, \alpha) = S_{\rm sd}(\alpha), \qquad (65)$$

the Fourier transform of the hybrid primary beam, $\overline{B}_{\rm hyb}$, as

$$\overline{B}_{\rm hyb}(u_p \neq 0, u') = \overline{B}(u_p - u') \qquad (66)$$
$$\overline{B}_{\rm hyb}(u_p = 0, u') = \overline{B}_{\rm sd}(-u'), \qquad (67)$$

and a hybrid weighting function, $W_{\rm hyb}$, as

$$W_{\rm hyb}(u_p \neq 0, u' + u'' - u_p) = W_{\rm hyb}(u_p, u' + u'' - u_p), \quad (68)$$
$$W_{\rm hyb}(u_p = 0, u' + u'') = W_{\rm sd}(u' + u''). \qquad (69)$$

All the processing steps described in the previous sections (including a potential gridding step of single-dish, on-the-fly data) can then be directly applied to the hybrid data set. Using the wide-field synthesis formalism, we can easily write

$$\overline{I}_{\rm hyb}(u) = \int_{u'} \overline{D}_{\rm hyb}(u', u - u') \, \overline{I}(u') \, {\rm d}u', \qquad (70)$$

with $\quad \overline{I}_{\rm hyb}(u) = \overline{I}_{\rm dirty}(u) + W_{\rm sd}(u) \, \overline{I}_{\rm sd}(u) \qquad (71)$

and $\overline{D}_{\rm hyb}(u', u'')$

$$= \overline{D}(u', u'') + W_{\rm sd}(u' + u'') \, \overline{S}_{\rm sd}(u'') \, \overline{B}_{\rm sd}(-u'). \qquad (72)$$

We thus see that $\overline{I}_{\rm hyb}$ is a linear combination of the information measured by the single-dish ($\overline{I}_{\rm sd}$) and by the interferometer ($\overline{I}_{\rm dirty}$). There, $W_{\rm sd}(u)$ plays a particular role for two reasons. First, its dependency on the spatial frequency (u) enables us to filter out the highest spatial frequencies that are measured by the single-dish antenna with low accuracy. Second, it is well-known that the relative weight of the single-dish to interferometric data is a critical parameter in the processing of the short spacings from single-dish data (see e.g. Rodríguez-Fernández et al. 2008). This relative weight is a free parameter within the restrictions set by the noise level (*i.e.*, we want the single-dish data to bring information and not just noise to the interferometric data), and a criterion must therefore be defined to adjust it to an optimal value. We refer the reader to the discussion of Sect. 5, which also applies here.

Table 5. Definition of the symbols used to expose the processing of the short spacings.

Symbol & Definition		Plane(s)
$I_{\rm sd}$	Measured single-dish intensity	sky
$B_{\rm sd}$	Single-dish antenna power pattern	sky
$S_{\rm sd}$	Single-dish sampling function	sky
$V_{\rm sd}$	Single-dish visibility function	sky
$W_{\rm sd}$	Single-dish uv-plane weighting function	uv
b_i	Voltage pattern of antenna i	sky
B_{ij}	Power pattern of antenna i and j ($= b_i \, b_j^\star$)	sky
V_{ij}	Visibility between antenna i and j	uv & sky
$I_{\rm hyb}$	Hybrid dirty image	sky
$B_{\rm hyb}$	Hybrid antenna power pattern	sky
$S_{\rm hyb}$	Hybrid sampling function	uv & sky
$W_{\rm hyb}$	Hybrid uv-plane weighting function	uv & uv
$D_{\rm hyb}$	Set of hybrid dirty beams	sky & sky

Notes. This table uses similar conventions as Table 1. The top part defines symbols related to single-dish measurements. The middle part defines symbols related to heterogeneous-array measurements. The bottom part defines hybrid symbols, which results from combinations of single-dish and heterogeneous-array measurements.

6.4. Processing short spacings from heterogeneous arrays

A heterogeneous array is an interferometer composed with antennas of different diameters. ALMA and CARMA are two such examples. The measurement equation for a heterogeneous array is

$$V_{ij}(u_p, \alpha_s) = \int_{\alpha_p} b_i(\alpha_p - \alpha_s) \, b_j^\star(\alpha_p - \alpha_s) \, I(\alpha_p) \, {\rm e}^{-i2\pi\alpha_p u_p} \, {\rm d}\alpha_p, (73)$$

where b_i and b_j are the voltage reception patterns of the antenna pair that forms the ij baseline and the asterisk denotes the complex conjugate (Thompson et al. 1986, chapter 3). The formalism developed in the previous sections holds as long as we redefine

$$B_{ij}(\alpha) \equiv b_i(\alpha) \, b_j^\star(\alpha). \qquad (74)$$

A simple application of the correlation theorem implies that

$$\overline{B}_{ij}(u) = \int_{u'} \overline{b}_i(u + u') \, \overline{b}_j(u') \, {\rm d}u'. \qquad (75)$$

The use of the baseline indices ij must be generalized throughout the equations because the knowledge of the antenna type must be attached to each individual data point (visibility). As a result, the wide-field synthesis formalism can be easily adapted to heterogeneous arrays at the price of additional bookkeeping.

6.5. Two textbook cases: IRAM-30m + PdBI and ALMA + ACA

Figure 5 sketches why wide-field synthesis naturally handles the short spacings in two textbook cases. In the ideal case, the Fourier transform along the α_s dimension produces visibilities, which are related to the wide-field spatial frequencies of the source brightness weighted by the transfer function of the interferometer. In this sense, Fig. 5 displays the natural weighting of the synthesized wide-field visibilities at the position of each measured visibility. Handling visibilities from antenna of different sizes just implies that
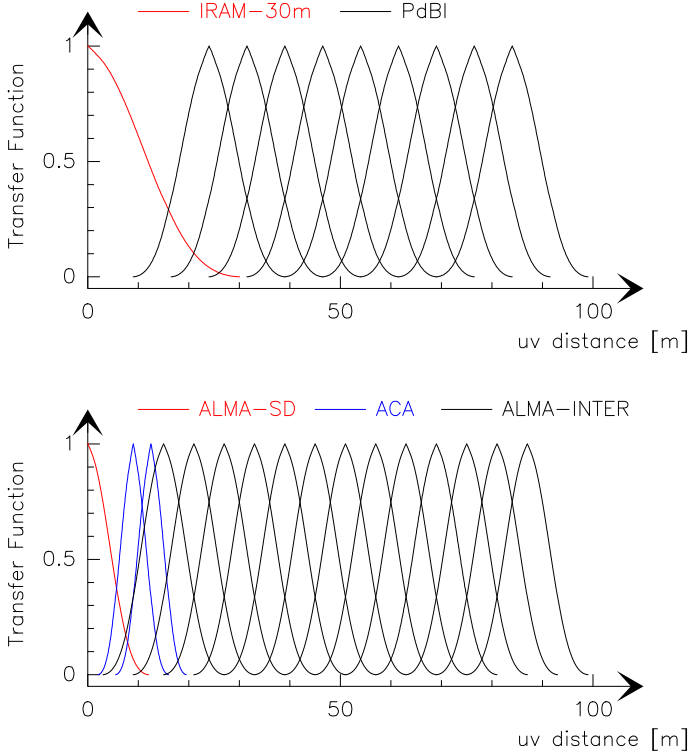
Fig. 5. Sketches of the natural weighting of the synthesized wide-field visibilities. Each measured spatial frequency will produce wide-field spatial frequencies apodized by the transfer function ($\overline{B}$) centered on the measured spatial frequency. The used transfer function depends on the telescopes used, explaining why wide-synthesis naturally handles the short spacing either from a single-dish antenna or from a heterogeneous array. The synthesized visibilities in the overlapping regions will then be averaged. Two textbook examples are illustrated: 1) the combination of data from the IRAM-30m single-dish (red transfer function) and from the Plateau de Bure Interferometer (black transfer functions) at the top; and 2) the combination of data from ALMA 12m-antennas used either in single-dish mode (red transfer function), in interferometric mode (black transfer functions) and of data from the ACA 7m-antennas (blue transfer functions) at the bottom. The minimum uv distances measured by each interferometer were set from the minimum possible distance between antennas (24 m for PdBI, 15 m for ALMA and 9 m for ACA).

the natural weighting function of the synthesized visibilities will have a different shape.

The top panel of Fig. 5 displays how the IRAM-30m single-dish is used to complement the Plateau de Bure interferometer visibilities. The bottom panel displays how ACA is used to produce the short spacing information for ALMA. The four 12m-antennas will provide the single-dish information, while the 12 additional 7m-antennas will form with ALMA a heterogeneous array. In the first design, ACA and ALMA form two independent interferometers; *i.e.*, they are not cross-correlated. The single-dish antennas, ACA and ALMA, thus appear as three different instruments. It is thus possible to decompose the hybrid set of wide-field dirty

beams obtained by processing the 3 sets of data together in 3 different sets of dirty beams

$$\overline{D}_{\rm hyb}(u', u'') = \overline{D}_{12\rm m}(u', u'') + \overline{D}_{7\rm m}(u', u'') + \overline{D}_{\rm sd}(u', u''), (76)$$

with $\overline{D}_{\rm hyb}(u', u'') \equiv$

$$\int_{u_p} W_{\rm hyb}(u_p, u' + u'' - u_p)\, \overline{S}_{\rm hyb}(u_p, u'')\, \overline{B}_{\rm hyb}(u_p, u')\, {\rm d}u_p. (77)$$

For a multiplying interferometer,

$$\forall\, |u_p| < d_{\rm prim}, \quad S(u_p, \alpha_s) = 0 \quad {\rm and} \quad \overline{S}(u_p, u_s) = 0. \quad (78)$$

This implies that $\overline{D}_{\rm sd}$ contributes at $u_p = 0$ in the sum over u_p in Eq. 77, $\overline{D}_{7\rm m}$ contributes for $9\,{\rm m} < u_p \lesssim 40\,{\rm m}$ and $\overline{D}_{12\rm m}(u', u'')$ contributes for $15\,{\rm m} < u_p < 150\,{\rm m}$ in the most compact configuration of ALMA.

7. Comparison with standard nonlinear mosaicking

7.1. Mosaicking in a nutshell

Several excellent descriptions of the mosaicking imaging and deconvolution algorithms can be found (see e.g. Cornwell 1988; Cornwell et al. 1993; Sault et al. 1996b). Here, we summarize the approach implemented in the GILDAS/MAPPING software used to image and deconvolve the data from the Plateau de Bure Interferometer. This approach is based on original ideas by F. Viallefond in the early 90s (Gueth et al. 1995).

The basic ideas of nonlinear mosaicking are 1) imaging the different fields of the mosaic independently, 2) linearly adding the single-field dirty images into a dirty mosaic, and 3) jointly deconvolving the dirty mosaic.

7.1.1. Single-field imaging

For simplicity, we skip the gridding convolution in the following equations because the gridding step does not change the nature of the equations. Imaging the fields individually means that we will work at constant α_s. We first define the single-field dirty image of the α_s-field as

$$I_{\rm sfd}(\alpha_p, \alpha_s) \underset{u_p}{\overset{\alpha_p}{\supset}} \overline{I}_{\rm sfd}(u_p, \alpha_s), \quad (79)$$

where the Fourier transform of the single-field dirty image is the product of the sampling function $S(u_p, \alpha_s)$ and the visibility function $V(u_p, \alpha_s)$:

$$\overline{I}_{\rm sfd}(u_p, \alpha_s) \equiv SV(u_p, \alpha_s). \quad (80)$$

From the previous equations, it is easily demonstrated that

$$I_{\rm sfd}(\alpha_p, \alpha_s) = \int_{\alpha_p'} \Delta(\alpha_p - \alpha_p', \alpha_s) \left[B(\alpha_p' - \alpha_s) I(\alpha_p') \right]\, {\rm d}\alpha_p', (81)$$

where the single-field dirty beam is defined as

$$\Delta(\alpha_p, \alpha_s) \underset{u_p}{\overset{\alpha_p}{\supset}} S(u_p, \alpha_s). \quad (82)$$

We can rewrite the previous equation as

$$I_{\rm sfd}^{\alpha_s}(\alpha_p) = \{ \Delta^{\alpha_s} \star (B^{\alpha_s} I) \}(\alpha_p), \quad (83)$$

meaning that the single-field dirty images can be written as a local convolution of $B^{\alpha_s} I$ and Δ^{α_s}, the single-field dirty beam associated to the currently imaged field.

Table 6. Definition of the symbols used to expose the mosaicking formalism.

Symbol & Definition		Plane(s)
α_s	Scanned angle	sky
u_s	Scanned spatial frequency	uv
α_p	Phased angle	sky
u_p	Phased spatial frequency	uv
I	Sky brightness	sky
B	Primary beam	sky
V	Visibility function	uv & sky
S	Sampling function	uv & sky
Δ	Set of single-field dirty beams	sky & sky
I_{sfd}	Set of single-field dirty images	sky & sky
D_{mos}	Set of mosaicked dirty beams	sky & sky
Ω_{mos}	Mosaicking sky weighting function	sky & sky
W_{mos}	Mosaicking uv weighting function	uv & uv
I_{mos}	Mosaicked dirty image	sky
N_{mos}	Mosaicked noise image	sky
R	Residual image	sky
SNR	Signal-to-noise ratio image	sky

Notes. This table uses similar conventions as Table 1. The top part repeats the symbols which have the same meaning in both the mosaicking and the wide-field synthesis formalisms, while the bottom part defines symbols that either have a different meaning in both formalisms or are used only in the mosaicking formalism.

7.1.2. Mosaicking the dirty images

In GILDAS/MAPPING[7], the single-field dirty images are formed on the same grid (in particular the same pixel size and the same image size covering about twice the mosaic field of view). These single-field dirty images are then linearly averaged as

$$I_{\mathrm{mos}}(\alpha_p) \equiv \int_{\alpha_s} \Omega_{\mathrm{mos}}(\alpha_p, \alpha_s)\, I_{\mathrm{sfd}}(\alpha_p, \alpha_s)\, \mathrm{d}\alpha_s, \qquad (84)$$

where $\quad \Omega_{\mathrm{mos}}(\alpha_p, \alpha_s) \equiv \dfrac{w(\alpha_s)\, B(\alpha_p - \alpha_s)}{\int_{\alpha_s} w(\alpha_s)\, B^2(\alpha_p - \alpha_s)\, \mathrm{d}\alpha_s} \qquad (85)$

and $w(\alpha_s)$ is the sky plane weighting function, i.e.,

$$w(\alpha_s) = \sum_i \delta(\alpha_s - \alpha_i) \frac{1}{\sigma_i^2}. \qquad (86)$$

In the previous equation, the α_i are the positions of each sky-plane measurement, and σ_i is the rms noise associated with $I_{\mathrm{sfd}}^{\alpha_i}$. Cornwell et al. (1993) demonstrates that the noise in the mosaic image naturally varies across the field as

$$N_{\mathrm{mos}}(\alpha_p) = \frac{1}{\sqrt{\int_{\alpha_s} w(\alpha_s)\, B^2(\alpha_p - \alpha_s)\, \mathrm{d}\alpha_s}}. \qquad (87)$$

In particular, it rises sharply at the edges of the mosaic.

7.1.3. Joint deconvolution

Standard algorithms of single-field deconvolution must be adapted to the mosaicking case because both the dirty

----

[7] See http://www.iram.fr/IRAMFR/GILDAS for more information about the GILDAS software.

----

beam and the noise vary across the mosaic field of view. We describe here the adaptations made in GILDAS/MAPPING of the simplest CLEAN deconvolution method, described in Högbom (1974). Adaptations of more evolved CLEAN deconvolution methods are also implemented following the same basic rules.

1. We first initialize the residual and signal-to-noise maps from the dirty and noise maps

$$R_0(\alpha_p) = I_{\mathrm{mos}}(\alpha_p) \qquad (88)$$

and $\quad \mathrm{SNR}_0(\alpha_p) = \dfrac{I_{\mathrm{mos}}(\alpha_p)}{N_{\mathrm{mos}}(\alpha_p)}. \qquad (89)$

2. The kth CLEAN component is sought on the SNR_{k-1} map instead of the R_{k-1} map to ensure that noise peaks at the edges of the mosaic are not confused with the true signal of the same magnitude.
3. Using that the kth CLEAN component is a point source of intensity I_k at position α_k, the residual and signal-to-noise maps are then upgraded as follows:
$$R_k(\alpha_p) = R_{k-1}(\alpha_p)$$
$$-\gamma I_k \int_{\alpha_s} \Omega_{\mathrm{mos}}(\alpha_p, \alpha_s)\, \Delta(\alpha_p - \alpha_k, \alpha_s)\, B(\alpha_k - \alpha_s)\, \mathrm{d}\alpha_s, \quad (90)$$

and $\quad \mathrm{SNR}_k(\alpha_p) = \dfrac{R_k(\alpha_p)}{N_{\mathrm{mos}}(\alpha_p)}. \qquad (91)$

Here $\gamma (\sim 0.2)$ is the usual loop gain that ensures convergence of the CLEAN algorithms.
4. Steps 2 and 3 are iterated as long as the stopping criterion is not met.

7.1.4. Wide-field measurement equation

To help the comparison between mosaicking and wide-field synthesis, we now go one step further than is usually done in the description of mosaicking; i.e., we write the image-plane measurement equation as a wide-field measurement equation of the same kind as Eq. 23. Substituting Eq. 81 into Eq. 84 and reordering the terms after inverting the order of the sum over α_s and α_p, one obtains

$$I_{\mathrm{mos}}(\alpha_p) = \int_{\alpha_p'} D_{\mathrm{mos}}(\alpha_p - \alpha_p', \alpha_p)\, I(\alpha_p')\, \mathrm{d}\alpha_p', \qquad (92)$$

with $D_{\mathrm{mos}}(\alpha', \alpha'')$

$$= \int_{\alpha_s} B(\alpha'' - \alpha' - \alpha_s)\, \Omega_{\mathrm{mos}}(\alpha'', \alpha_s)\, \Delta(\alpha', \alpha_s)\, \mathrm{d}\alpha_s. \qquad (93)$$

Taking the inverse Fourier transforms of D_{mos}, we get the mosaicking transfer function
$$\overline{D}_{\mathrm{mos}}(u', u - u') =$$

$$\iint_{u_p u_s} W_{\mathrm{mos}}(u - u_p, u_s - u')\, \overline{S}(u_p, u_p - u_s)\, \overline{B}(u_p - u')\, \mathrm{d}u_p \mathrm{d}u_s, (94)$$

with $\quad \Omega_{\mathrm{mos}}(\alpha', \alpha'') \overset{(\alpha', \alpha'')}{\underset{(u', u'')}{\supset}} W_{\mathrm{mos}}(u', u''). \qquad (95)$

7.2. Comparison

While both mosaicking and wide-field synthesis produce image-plane measurement equations of the same kind (see Eqs. 23 and 92), the comparison of the dirty beams (Eqs. 57 and 93) and of the transfer functions (Eqs. 35 and 94) immediately shows the different dependencies on the primary beams (B), the single-field dirty beams (Δ), the image-plane weighting functions (Ω), and their respective Fourier transforms ($\overline{B}$, $\overline{S}$ and W). This means that mosaicking is not mathematically equivalent to wide-field synthesis, though both methods recover the sky brightness. These differences come directly from the differences in the processing. If we momentarily forget the gridding steps, mosaicking starts with a Fourier transform along the u_p dimension of the visibility function, and most of the processing thus happens in the sky plane, while wide-field synthesis starts with a Fourier transform along the α_s dimension, and most of the processing thus happens in the uv plane.

Moreover, both methods are irreducible to each other. Wide-field synthesis gives a more complex dirty beam formulation in the image plane, which could give the impression that it is a generalization of mosaicking. Indeed, the wide-field image-plane weighting function can be chosen as the product of a Dirac function of α' times a function ω of α'',

$$i.e., \quad \Omega(\alpha', \alpha'') = \delta(\alpha')\, \omega(\alpha''). \qquad (96)$$

This implies a wide-field uv-plane weighting function independent of u'; *i.e.*, $W(u', u'') = \overline{\omega}(u'')$. This choice is a clear limitation because it enables us to influence the transfer function only locally (around each measured u_p spatial frequency), while weighting is generally intended to globally influence the transfer function (see Sect. 5). Eitherway, in this case, the wide-field dirty beam can easily be simplified to

$$D(\alpha', \alpha'') = \int_{\alpha_s} B(\alpha'' - \alpha' - \alpha_s)\, \omega(\alpha'' - \alpha_s)\, \Delta(\alpha', \alpha_s)\, d\alpha_s. \quad (97)$$

While this simplified formulation of the wide-field dirty beam is closer to the mosaicking formulation, they still differ in a major way: $\omega(\alpha'' - \alpha_s)$ is a shift-invariant function contrary to $\Omega_{\mathrm{mos}}(\alpha'', \alpha_s)$. This is the shift-dependent property of $\Omega_{\mathrm{mos}}(\alpha'', \alpha_s)$, which implies the additional complexity (integral over u_s in addition to the integral over u_p) of the mosaicking (Eq. 94) over the wide-field transfer function (Eq. 35).

One main difference between the two processing methods is that standard mosaicking prescribes a precise weighting function, while we argue that the wide-field weighting function should be defined according to the context (see Sect. 5). Another important difference is the treatment of the short spacings, which are naturally processed in the wide-field synthesis methods, but which needs a very specific treatment in mosaicking (see Sect. 6 and references therein). Finally, while mosaicking implies a gridding only of u_p dimension of the measured visibilities, wide-field synthesis naturally requires a gridding of both the u_p and α_s dimensions. As the Nyquist sampling along the α_s dimension is only $0.5/d_{\mathrm{prim}}$, the gridding of the sky plane can result in a large reduction of the data storage space and cpu processing cost when processing on-the-fly and/or multi-beam observations.

8. Summary

Interferometric wide-field imaging implies scanning the sky in one way or another (e.g. stop-and-go mosaicking, on-the-fly scanning, sampling of the focal plane by multi-beams). This produces sampled visibilities SV, which depends both on the uv-plane and sky coordinates (e.g., u_p and α_s).

Based on a basic idea by Ekers & Rots (1979), we proposed a new way to image the interferometric wide-field sampled visibilities: $SV(u_p, \alpha_s)$. After gridding the measured visibilities both in the uv and sky planes, the gridded visibilities $SV^{\mathcal{G}}$ are Fourier-transformed along the α_s sky dimension, yielding synthesized visibilities $\overline{SV}^{\mathcal{G}}$ sampled on a uv grid whose cell size is related to the total field of view; *i.e.*, it is much finer than the diameter of the interferometer antennas. We thus proposed calling this processing scheme "wide-field synthesis".

The Fourier transform is performed for each constant u_p value. As many independent estimates of the uv plane are produced as independent values of u_p measured. A shift-and-average operator is then used to build a final, wide-field uv plane, which translates into a wide-field dirty image after inverse Fourier transform, *i.e.*,

$$\overline{I}_{\mathrm{dirty}}^{\mathcal{G}}(u) \equiv \int_{u_p} W(u_p, u - u_p)\, \overline{SV}^{\mathcal{G}}(u_p, u - u_p)\, du_p, \qquad (98)$$

where W is a normalized weighting function. Using these tools, we demonstrated that:

1. The dirty image ($I_{\mathrm{dirty}}^{\mathcal{G}}$) is a convolution of the sky brightness distribution (I) with a set of wide-field dirty beams ($D^{\mathcal{G}}$) varying with the sky coordinate α, *i.e.*,

$$I_{\mathrm{dirty}}^{\mathcal{G}}(\alpha) = \int_{\alpha'} D^{\mathcal{G}}(\alpha - \alpha', \alpha)\, I(\alpha')\, d\alpha'. \qquad (99)$$

Compared to single-field imaging, the dependency on the primary beam is transferred from a product of the sky brightness distribution into the definition of the set of wide-field dirty beams.

2. The set of gridded dirty beams ($D^{\mathcal{G}}$) can be computed from the ungridded sampling function (S), the transfer function ($\overline{B}$, the inverse Fourier transform of the primary beam), and the gridding convolution kernel (See Eq. 42, 50 and 51).

3. The dependence of the wide-field dirty beams on the sky position is slowly-varying, with their shape varying on an angular scale typically larger than or equal to the primary beamwidth.

Adaptations of the existing deconvolution algorithms should be straightforward.

A comparison with standard nonlinear mosaicking shows that it is not mathematically equivalent to the wide-field synthesis proposed here, though both methods do recover the sky brightness. The main advantages of wide-field synthesis over standard nonlinear mosaicking are

1. Weighting is at the heart of the wide-field synthesis because it is an essential part of the shift-and-average operation. Indeed, not only can a multiplicative weight be attributed to each visibility sample before any processing, but the uv-plane weighting function (W, see Eq. 98) is also a degree of freedom, which should be set according to the conditions of the observation and the imaging

goals, e.g. highest signal-to-noise ratio, highest resolution, or most uniform resolution over the field of view. The W weighting function thus enables us to modify the wide-field response of the instrument. On the other hand, mosaicking requires a precise weighting function in the image plane, which freezes the wide-field response of the interferometer.

2. Wide-field synthesis naturally processes the short spacings from both single-dish antennas and heterogeneous arrays along with the long spacings. Both of them can then be jointly deconvolved.

3. The gridding of the sky plane dimension of the measured visibilities, required by the wide-field synthesis, may potentially save large amounts of hard-disk space and cpu processing power relative to mosaicking when handling data sets acquired with the on-the-fly observing mode. Wide-field synthesis could thus be particularly well suited to process on-the-fly observations.

The wide-field synthesis algorithm is compatible with the uvw-unfaceting technique devised by Sault et al. (1996a) to deal with the celestial projection effect, known as non-coplanar baselines (see appendix B). Finally, on-the-fly observations imply an elongation of the primary beam along the scanning direction. These effects can be decreased by an increase in the primary beam sampling rate. However, it may limit the dynamic range of the image brightness if the primary beam sampling rate is too coarse (see appendix C).

References

Bevington, P. R. & Robinson, D. K. 2003, Data Reduction and Error Analysis For the Physical Sciences, 3rd edn. (McGraw-Hill)

Bhatnagar, S. & Cornwell, T. J. 2004, A&A, 426, 747

Bhatnagar, S., Cornwell, T. J., Golap, K., & Uson, J. M. 2008, A&A, 487, 419

Bracewell, R. N. 2000, The Fourier Transform and its Applications, 3rd edn. (McGraw-Hill)

Clark, B. G. 1980, A&A, 89, 377

Cornwell, T. J. 1988, A&A, 202, 316

Cornwell, T. J., Golap, K., & Bhatnagar, S. 2008, IEEE Journal of Selected Topics in Signal Processing, Vol. 2, Issue 5, p.647-657, 2, 647

Cornwell, T. J., Holdaway, M. A., & Uson, J. M. 1993, A&A, 271, 697

Cotton, W. D. & Uson, J. M. 2008, A&A, 490, 455

D'Addario, L. R. & Emerson, D. T. 2000, On-The-Fly Fringe Tracking, ALMA memo 331

Ekers, R. D. & Rots, A. H. 1979, in IAU Proc., Vol. 49, Image Formation from Coherence Fucntions in Astronomy, ed. C. van Schoonedveld (D. Reidel), p. 61

Frater, R. H. & Docherty, I. S. 1980, A&A, 84, 75

Goldsmith, P. F. 1998, Gaussian Beam, Quasioptical Propagation and Applications (IEEE Press)

Gueth, F., Guilloteau, S., & Viallefond, F. 1995, in The XXVIIth Young European Radio Astronomers Conference, ed. D. A. Green & W. Steffen, 8–+

Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, A&AS, 117, 137

Högbom, J. A. 1974, A&AS, 15, 417

Holdaway, M. A. & Foster, S. M. 1994, On-The-Fly Mosaicing, ALMA memo 122

Mangum, J. G., Emerson, D. T., & Greisen, E. W. 2007, A&A, 474, 679

Narayan, R. & Nityananda, R. 1986, ARA&A, 24, 127

Pety, J., Gueth, F., & Guilloteau, S. 2001, Impact of ACA on the Wide-Field Imaging Capabilities of ALMA, ALMA memo 398

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical Recipes in C, 2nd edn. (Cambridge University Press)

Rodríguez-Fernández, N. J., Pety, J., & Gueth, F. 2008, Single-dish observation and processing to produce the short-spacing information for a millimeter interferometer, IRAM memo 2008-2

Rodríguez-Fernández, N. J., Pety, J., & Gueth, F. 2009, Imaging of interferometic On-The-Fly observations: Context and discussion of possible methods, IRAM memo 2009-2

Sault, R. J., Bock, D. C.-J., & Duncan, A. R. 1999, A&AS, 139, 387

Sault, R. J., Hamaker, J. P., & Bregman, J. D. 1996a, A&AS, 117, 149

Sault, R. J., Staveley-Smith, L., & Brouw, W. N. 1996b, A&AS, 120, 375

Schwab, F. R. 1984, AJ, 89, 1076

Sramek, R. A. & Schwab, F. R. 1989, Synthesis Imaging in Radio Astronomy, Conference series (Astronomical Society of the Pacific), 117–

Thompson, A. R., Moran, J. M., & Swenson, G. W. J. 1986, Interferometry and Synthesis in Radio Astronomy (John Wiley & Sons)

Appendix A: Demonstrations

A.1. Ekers & Rots scheme

Fourier-transforming the visibility function along the α_s dimension at constant u_p, we derive with simple replacements

$$\overline{V}_{u_p}(u_s) \tag{A.1}$$

$$= \int_{\alpha_s} V_{u_p}(\alpha_s)\, e^{-i2\pi\alpha_s u_s}\, d\alpha_s \tag{A.2}$$

$$= \iint_{\alpha_s\alpha_p} B(\alpha_p - \alpha_s)\, I(\alpha_p)\, e^{-i2\pi(\alpha_p u_p + \alpha_s u_s)}\, d\alpha_s d\alpha_p. \tag{A.3}$$

We then use the following change of variables $\beta \equiv \alpha_p - \alpha_s$ and $d\beta = -d\alpha_s$, to get

$$\overline{V}_{u_p}(u_s) \tag{A.4}$$

$$= \iint_{\alpha_p\beta} B(\beta)\, I(\alpha_p)\, e^{-i2\pi\left[\alpha_p(u_p + u_s) - \beta u_s\right]}\, d\alpha_p d\beta \tag{A.5}$$

$$= \left[\int_\beta B(\beta)\, e^{-i2\pi\beta(-u_s)}\, d\beta\right]\left[\int_{\alpha_p} I(\alpha_p)\, e^{-i2\pi\alpha_p(u_p + u_s)}\, d\alpha_p\right] \tag{A.6}$$

$$= \overline{B}(-u_s)\, \overline{I}(u_p + u_s). \tag{A.7}$$

A.2. Incomplete sampling

We here demonstrate that Eq. 23 and Eq. 34 are equivalent. To do this, we take the direct Fourier transform of $I_{\mathrm{dirty}}(\alpha)$

$$\overline{I}_{\mathrm{dirty}}(u) \tag{A.8}$$

$$= \iint_{\alpha\alpha'} D(\alpha - \alpha', \alpha)\, I(\alpha')e^{-i2\pi\alpha u}\, d\alpha d\alpha', \tag{A.9}$$

and we replace $I(\alpha')$ by its formulation as a function of its Fourier transform

$$I(\alpha') = \int_{u'} \overline{I}(u')\, e^{+i2\pi u'\alpha'}\, du'. \tag{A.10}$$

We thus derive $\overline{I}_{\mathrm{dirty}}(u)$

$$= \int_{u'} \left[\iint_{\alpha\alpha'} D(\alpha - \alpha', \alpha)e^{-i2\pi(\alpha u - \alpha' u')}\, d\alpha d\alpha'\right]\overline{I}(u')\, du'. \tag{A.11}$$

Using the following change of variables $\alpha'' \equiv \alpha - \alpha'$, $\alpha' = \alpha - \alpha''$ and $d\alpha'' = -d\alpha'$, the innermost integral can be written as

$$\iint_{\alpha\alpha'} D(\alpha - \alpha', \alpha)\, e^{-i2\pi(\alpha u - \alpha' u')}\, d\alpha d\alpha' \tag{A.12}$$

$$= \int_\alpha \left[\int_{\alpha''} D(\alpha'', \alpha)\, e^{-i2\pi\alpha'' u'}\, d\alpha''\right] e^{-i2\pi\alpha(u - u')}\, d\alpha \tag{A.13}$$

$$= \int_\alpha \underline{D}(u', \alpha)\, e^{-i2\pi\alpha(u - u')}\, d\alpha \tag{A.14}$$

$$= \overline{\underline{D}}(u', u - u'). \tag{A.15}$$

In the last two steps, we have simply recognized two different steps of Fourier transforms of D. Finally,

$$\overline{I}_{\mathrm{dirty}}(u) = \int_{u'} \overline{\underline{D}}(u', u - u')\, \overline{I}(u')\, du'. \tag{A.16}$$

A.3. Gridding

The gridding kernel can be defined as the product of two functions, each one operating in its own dimension. We use this to study separately the effect of gridding in the uv and sky planes. We then use the intermediate results to get the effect of gridding simultaneously in both planes.

A.4. Gridding in the uv plane

We define the sampled visibility function gridded in the uv plane as

$$SV^g(u_p, \alpha_s) \equiv \{g \star SV^{\alpha_s}\}(u_p) \tag{A.17}$$

$$= \int_{u'_p} g(u_p - u'_p)\, SV^{\alpha_s}(u'_p)\, du'_p, \tag{A.18}$$

Using that the gridding is here applied on the u_p dimension, while the Fourier transform is applied on the α_s dimension, it is easy to show that the gridding and Fourier-transform operations commute:

$$\overline{SV}^g_{u_p}(u_s) \tag{A.19}$$

$$= \iint_{\alpha_s u'_p} g(u_p - u'_p)S(u'_p, \alpha_s)V(u'_p, \alpha_s)e^{-i2\pi\alpha_s u_s}\, d\alpha_s du'_p \tag{A.20}$$

$$= \int_{u'_p} g(u_p - u'_p)\, \overline{SV}_{u'_p}(u_s)\, du'_p \tag{A.21}$$

Defining the Fourier transform of the uv gridded dirty image, we derive

$$\overline{I}^g_{\mathrm{dirty}}(u) \equiv \left\langle \overline{SV}^g \right\rangle(u) \tag{A.22}$$

$$= \iint_{u_p u'_p} W(u_p, u - u_p)\, g(u_p - u'_p)\, \overline{SV}_{u'_p}(u - u_p)\, du_p du'_p \tag{A.23}$$

Using Eq. 31 to replace $\overline{SV}_{u'_p}(u - u_p)$, we can write the Fourier transform of the uv gridded dirty image as

$$\overline{I}^g_{\mathrm{dirty}}(u) = \int_{u'} \overline{\underline{D}}^g(u', u - u')\, \overline{I}(u')\, du', \tag{A.24}$$

with

$$\overline{\underline{D}}^g(u', u - u') \equiv \int_{u_p} W(u_p, u - u_p)\, \underline{\Sigma}^g(u_p, u - u_p, u')\, du_p \tag{A.25}$$

and $\underline{\Sigma}^g(u_p, u_s, u')$

$$\equiv \int_{u'_p} g(u_p - u'_p)\, \overline{S}(u'_p, u_s - u' + u'_p)\, \overline{B}(u'_p - u')\, du'_p. \tag{A.26}$$

Using $\overline{S}(u'_p, u_s - u' + u'_p)$

$$= \left[\int_{\alpha_s} S_{u'_p}(\alpha_s)\, e^{-i2\pi\alpha_s u_s}\, d\alpha_s\right] e^{-i2\pi(u'_p - u')\alpha_s}, \tag{A.27}$$

and $\overline{\Sigma}^g(u_p, \alpha_s, u') \overset{\alpha_s}{\underset{u_s}{\supset}} \underline{\Sigma}^g(u_p, u_s, u')$, $\tag{A.28}$

we derive $\overline{\Sigma}^g(u_p, \alpha_s, u') =$

$$\int_{u'_p} g(u_p - u'_p)\, S(u'_p, \alpha_s)\, \overline{B}(u'_p - u')\, e^{-i2\pi(u'_p - u')\alpha_s}\, du'_p, \tag{A.29}$$

or

$$\overline{\Sigma}^g(u_p, \alpha_s, u') = \int_{u'_p} g(u_p - u'_p)\,\overline{\Sigma}(u'_p, \alpha_s, u'_p - u')\,\mathrm{d}u'_p. \quad \text{(A.30)}$$

Thus, $\overline{\Sigma}^g$ is the uv gridded version of the generalized sampling function $\overline{\Sigma}$.

A.4.1. Gridding in the sky plane

We define the sampled visibility function gridded in the sky plane as

$$SV^\gamma(u_p, \alpha_s) \equiv \left\{ \gamma \star SV_{u_p} \right\}(\alpha_s) \quad \text{(A.31)}$$

$$= \int_{\alpha'_s} \gamma(\alpha_s - \alpha'_s)\,SV_{u_p}(\alpha'_s)\,\mathrm{d}\alpha'_s, \quad \text{(A.32)}$$

Applying the convolution theorem on the Fourier transform along the α_s dimension, we derive

$$\overline{SV}^\gamma_{u_p}(u_s) = \overline{\gamma}(u_s)\,\overline{SV}_{u_p}(u_s). \quad \text{(A.33)}$$

Defining the Fourier transform of the sky-gridded dirty image, we derive

$$\overline{I}^\gamma_{\mathrm{dirty}}(u) \equiv \left\langle \overline{SV}^\gamma \right\rangle(u) \quad \text{(A.34)}$$

$$= \int_{u_p} W(u_p, u - u_p)\,\overline{\gamma}(u - u_p)\,\overline{SV}_{u_p}(u - u_p)\,\mathrm{d}u_p. \quad \text{(A.35)}$$

Using Eq. 31 to replace $\overline{SV}_{u_p}(u - u_p)$, we can write the Fourier transform of the sky-gridded dirty image as

$$\overline{I}^\gamma_{\mathrm{dirty}}(u) = \int_{u'} \overline{\underline{D}}^\gamma(u', u - u')\,\overline{I}(u')\,\mathrm{d}u' \quad \text{(A.36)}$$

with $\overline{\underline{D}}^\gamma(u', u - u')$

$$\equiv \int_{u_p} W(u_p, u - u_p)\,\overline{\underline{\Sigma}}^\gamma(u_p, u - u_p, u_p - u')\,\mathrm{d}u_p \quad \text{(A.37)}$$

and $\quad \overline{\underline{\Sigma}}^\gamma(u_p, u_s, u''_s) \equiv \overline{\gamma}(u_s)\,\overline{S}(u_p, u_s + u''_s)\,\overline{B}(u''_s), \quad \text{(A.38)}$

or, with the definition of $\overline{\underline{\Sigma}}$ (i.e., Eq. 45),

$$\overline{\underline{\Sigma}}^\gamma(u_p, u_s, u''_s) \equiv \overline{\gamma}(u_s)\,\overline{\underline{\Sigma}}(u_p, u_s, u''_s). \quad \text{(A.39)}$$

Using $\quad \overline{\Sigma}^\gamma(u_p, \alpha_s, u''_s) \underset{u_s}{\overset{\alpha_s}{\supset}} \overline{\underline{\Sigma}}^\gamma(u_p, u_s, u''_s), \quad \text{(A.40)}$

and the convolution theorem when taking the inverse Fourier transform of $\overline{\underline{\Sigma}}^\gamma$, we derive

$$\overline{\Sigma}^\gamma(u_p, \alpha_s, u''_s) = \int_{\alpha'_s} \gamma(\alpha_s - \alpha'_s)\,\overline{\Sigma}(u_p, \alpha'_s, u''_s)\,\mathrm{d}\alpha'_s. \quad \text{(A.41)}$$

Thus, $\overline{\Sigma}^\gamma$ is the sky gridded version of the generalized sampling function $\overline{\Sigma}$.

A.5. Gridding in both planes

Starting from the definition of $SV^\mathcal{G}$ (Eq. 41), we Fourier-transform it along the sky dimension at constant u_p. Using that the gridding along the u_p dimension can be factored out of the Fourier transform, we derive

$$\overline{SV}^\mathcal{G}_{u_p}(u_s) = \int_{u'_p} g(u_p - u'_p)\,\overline{SV}^\gamma_{u'_p}(u_s)\,\mathrm{d}u'_p. \quad \text{(A.42)}$$

Using Eq. A.33, we now replace $\overline{SV}^\gamma_{u'_p}(u_s)$ in the previous equation to get

$$\overline{SV}^\mathcal{G}_{u_p}(u_s) = \overline{\gamma}(u_s) \int_{u'_p} g(u_p - u'_p)\,\overline{SV}_{u'_p}(u_s)\,\mathrm{d}u'_p, \quad \text{(A.43)}$$

or $\quad \overline{SV}^\mathcal{G}_{u_p}(u_s) = \overline{\gamma}(u_s)\overline{SV}^g_{u_p}(u_s). \quad \text{(A.44)}$

From this relation, it is easy to deduce that

$$\overline{\underline{\Sigma}}^\mathcal{G}(u_p, u_s, u') = \overline{\gamma}(u_s)\overline{\underline{\Sigma}}^g(u_p, u_s, u'). \quad \text{(A.45)}$$

Using the convolution theorem when taking the inverse Fourier transform of $\overline{\underline{\Sigma}}^\mathcal{G}$ along the u_s dimension and replacing $\overline{\Sigma}^g(u_p, \alpha'_s, u')$ with Eq. A.30, we finally derive $\overline{\Sigma}^\mathcal{G}(u_p, \alpha_s, u')$

$$\equiv \iint_{u'_p \alpha'_s} g(u_p - u'_p)\,\gamma(\alpha_s - \alpha'_s)\,\overline{\Sigma}(u'_p, \alpha'_s, u'_p - u')\,\mathrm{d}u'_p \mathrm{d}\alpha'_s. \quad \text{(A.46)}$$

A.6. Wide-field vs single-field dirty beams

The notation 59 yields $W(u', u'') = \overline{\Omega}(u', u'')$. Using this in Eq. 35 gives $\overline{D}(u', u'')$

$$= \int_{u_p} \overline{\Omega}(u_p, u' + u'' - u_p)\,\overline{S}(u_p, u'')\,\overline{B}(u_p - u')\,\mathrm{d}u_p. \quad \text{(A.47)}$$

Taking the inverse Fourier transform along the u'' axis of Eq. A.47 and reordering the integral to factor out the term independent of u'', we can write

$$\underline{D}(u', \alpha'') = \int_{u_p} \overline{B}(u_p - u')\,\mathrm{FT}_1(u_p, u', \alpha'')\,\mathrm{d}u_p, \quad \text{(A.48)}$$

where

$$\mathrm{FT}_1(u_p, u', \alpha'') \quad \text{(A.49)}$$

$$\equiv \int_{u''} \overline{\Omega}(u_p, u' + u'' - u_p)\,\overline{S}(u_p, u'')\,\mathrm{e}^{+i2\pi u''\alpha''}\,\mathrm{d}u''. \quad \text{(A.50)}$$

We now introduce the following definition

$$\overline{S}(u_p, u'') \equiv \int_{\alpha_s} S(u_p, \alpha_s)\,\mathrm{e}^{-i2\pi\alpha_s u''}\,\mathrm{d}\alpha_s, \quad \text{(A.51)}$$

to derive

$$\mathrm{FT}_1(u_p, u', \alpha'')$$

$$= \int_{\alpha_s} S(u_p, \alpha_s)\left[\int_{u''} \overline{\Omega}(u_p, u' + u'' - u_p)\mathrm{e}^{+i2\pi u''(\alpha'' - \alpha_s)}\,\mathrm{d}u''\right]\mathrm{d}\alpha_s.$$

Using the following change of variables $v \equiv u'' + u' - u_p$, $u'' = v - u' + u_p$ and $dv = du''$ on the innermost integral, we get

$$\begin{aligned}
&\mathrm{FT}_1(u_p, u', \alpha'') \\
&= \int_{\alpha_s} S(u_p, \alpha_s)\, \underline{\Omega}(u_p, \alpha'' - \alpha_s)\, \mathrm{e}^{+i2\pi(u_p - u')(\alpha'' - \alpha_s)}\, \mathrm{d}\alpha_s.
\end{aligned}$$

Substituting this result into Eq. A.48 and taking the inverse Fourier transform along the u' axis, we can write
$$D(\alpha', \alpha'') =$$

$$\iint_{u_p \alpha_s} \underline{\Omega}(u_p, \alpha'' - \alpha_s)\, S(u_p, \alpha_s)\, \mathrm{FT}_2(u_p, \alpha_s, \alpha', \alpha'')\, \mathrm{d}u_p \mathrm{d}\alpha_s, \quad (\mathrm{A.52})$$

where

$$\begin{aligned}
&\mathrm{FT}_2(u_p, \alpha_s, \alpha', \alpha'') \\
&\equiv \int_{u'} \overline{B}(u_p - u')\, \mathrm{e}^{+i2\pi u_p(\alpha'' - \alpha_s)}\, \mathrm{e}^{+i2\pi u'(\alpha' - \alpha'' + \alpha_s)}\, \mathrm{d}u'.
\end{aligned}$$

Using the following change of variables $v \equiv u_p - u'$, $u' = u_p - v$ and $dv = du'$, we get

$$\mathrm{FT}_2(u_p, \alpha_s, \alpha', \alpha'') = B(\alpha'' - \alpha' - \alpha_s)\, \mathrm{e}^{+i2\pi u_p \alpha'}. \quad (\mathrm{A.53})$$

Substituting this result into Eq. A.52 and re-ordering the terms, we can write

$$D(\alpha', \alpha'') = \int_{\alpha_s} B(\alpha'' - \alpha' - \alpha_s)\, \mathrm{FT}_3(\alpha_s, \alpha', \alpha'')\, \mathrm{d}\alpha_s, \quad (\mathrm{A.54})$$

where

$$\mathrm{FT}_3(\alpha_s, \alpha', \alpha'') \equiv \int_{u_p} \underline{\Omega}(u_p, \alpha'' - \alpha_s)\, S(u_p, \alpha_s)\, \mathrm{e}^{+i2\pi u_p \alpha'}\, \mathrm{d}u_p.$$

A simple application of the convolution theorem gives

$$\mathrm{FT}_3(\alpha_s, \alpha', \alpha'') \equiv \int_{\alpha_p} \Omega(\alpha' - \alpha_p, \alpha'' - \alpha_s)\, \Delta(\alpha_p, \alpha_s)\, \mathrm{d}\alpha_p,$$

$$\text{where} \quad \Delta(\alpha_p, \alpha_s) \overset{\alpha_p}{\underset{u_p}{\supset}} S(u_p, \alpha_s). \quad (\mathrm{A.55})$$

Substituting this result into Eq. A.54, we finally derive the desired expression, *i.e.*, Eq. 57.

Appendix B: From the celestial sphere onto a single tangent plane

Eq. 1 neglects projection effects, known as non-coplanar baselines. Any method which deals with interferometric wide-field imaging must take this problem into account. After a short introduction to the problem, we show how wide-field synthesis is compatible with at least one method, namely the *uvw*-unfaceting of Sault et al. (1996b). This method tries to build a final wide-field *uv* plane from different pieces, just as our wide-field synthesis approach does. Another promising method is the *w*-projection, based on original ideas of Frater & Docherty (1980) and first successfully implemented by Cornwell et al. (2008). We did not look yet at its compatibility with wide-field synthesis.

B.1. *w-axis distortion*

When projection effects are taken into account, the measurement equation reads
$$V(w, u_p, \alpha_s)$$

$$= \int_{\alpha_p} B(\alpha_p - \alpha_s)\, \frac{I(\alpha_p)}{\sqrt{1 - \alpha_p^2}}\, \mathrm{e}^{-i2\pi\left[\alpha_p u_p + w(\sqrt{1 - \alpha_p^2} - 1)\right]}\, \mathrm{d}\alpha_p. (\mathrm{B.1})$$

In this equation, we continue to work in 1 dimension for the sky cosine direction (α_p), but we explicitly introduce the dependence along the direction perpendicular to the sky plane. This dependence appears in two ways, which is handled in very different ways. First, the factor $\sqrt{1 - \alpha_p^2}$ can be absorbed into a generalized sky brightness function

$$\mathcal{I}(\alpha_p) \equiv \frac{I(\alpha_p)}{\sqrt{1 - \alpha_p^2}}. \quad (\mathrm{B.2})$$

After imaging and deconvolution, $I(\alpha_p)$ can be easily restored from the deconvolved $\mathcal{I}(\alpha_p)$ image. The second dependence appears as an additional phase, which is written as

$$P(\alpha_p, w) \equiv \mathrm{e}^{-i2\pi w(\sqrt{1 - \alpha_p^2} - 1)}. \quad (\mathrm{B.3})$$

Thompson et al. (1986, chapter 4) shows that this additional phase can be neglected only if[8]

$$\frac{\pi}{4} \frac{\theta_{\mathrm{field}}^2}{\theta_{\mathrm{syn}}} \ll 1 \quad \text{or} \quad \pi \frac{\lambda d_{\mathrm{max}}}{d_{\mathrm{field}}^2} \ll 1. \quad (\mathrm{B.4})$$

The first form of the criterion indicates that the approximation gets worse at high spatial dynamic range (*i.e.*, $\theta_{\mathrm{field}}/\theta_{\mathrm{syn}} \ll 1$) while the second form indicates that the approximation gets worse at long wavelengths.

B.2. *uvw-unfaceting*

For stop-and-go mosaicking, it is usual to delay-track at the center of the primary beam for each pointing/field of the mosaic. This phase center is also the natural center of projection of each pointing/field. Stop-and-go mosaicking thus naturally paves the celestial sphere with as many tangent planes as there are pointings/fields; *i.e.*, this observing scheme is somehow enforcing a *uvw*-faceting scheme. In the framework of on-the-fly observations with ALMA, D'Addario & Emerson (2000) indicate that the phase center will be modified between each on-the-fly scan while it will stay constant during each on-the-fly scan. This is a compromise between loss of coherence and technical possibilities of the phase-locked loop. Using this hypothesis, the maximum sky area covered by the on-the-fly scan must take into account the maximum tolerable *w*-axis distortion.

The easiest way to deal with such data is to image each pointing/field around its phase center and then to reproject this image onto the mosaic tangent plane as displayed on Fig. 5 of Sault et al. (1996b). These authors point out that this scheme implies a typical *w*-axis distortion ϵ less than

$$\epsilon \leq (1 - \cos\theta_{\mathrm{alias}})\, \sin\theta_{\mathrm{center}} \sim \frac{1}{2}\, \theta_{\mathrm{center}}\, \theta_{\mathrm{alias}}^2, \quad (\mathrm{B.5})$$

---

[8] In contrast to the convention used in this paper, the d_{max} and d_{field} unit is meter instead of unit of λ in the second form of the criterion, in order to explicitly show the dependence on the wavelength.

where θ_{center} is the angle from the pointing/field center and θ_{alias} is the anti-aliasing scale defined in Sect. 4.2. In particular, ϵ is 0 at the phase center of each pointing/field. In other words, this scheme limits the magnitude of the w-axis distortion to its magnitude on a size equal to the anti-aliasing scale (*i.e.*, a few time the primary beamwidth) instead of a size equal to the total mosaic field of view. This scheme thus solves the projection effect as long as the w-axis distortion is negligible at sizes smaller than or equal to the anti-aliasing scale. A natural name for this processing scheme is uvw-unfaceting because it is the combination of a faceting observing mode (*i.e.*, regular change of phase center) and a linear transform of the uv coordinates to derive a single sine projection for the whole field of view.

Sault et al. (1996b) also demonstrate that the reprojection may be done much more easily and quickly in the uvw space before imaging the visibilities because it is then just a simple transformation of the uv coordinates, followed by a multiplication of the visibilities by a phase term. Finally, Sault et al. (1996b) note that it is the linear character of this uv coordinate transform which preserves the measurement equation 1. As the change of coordinates happens before any other processing, it also conserves all the equations derived in the previous sections to implement the wide-field synthesis.

Appendix C: On-the-fly observing mode and effective primary beam

Usual interferometric observing modes (including stop-and-go mosaicking) implies that the interferometer antennas observe a fixed point of the sky during the integration time. Conversely, the on-the-fly observing mode implies that the antennas slew on the sky during the integration time. This implies that the measurement equation 1 must be written as (Holdaway & Foster 1994; Rodríguez-Fernández et al. 2009):

$$V(\hat{u}_p, \hat{\alpha}_s) =$$

$$\frac{1}{\delta t} \int_{t_0-\delta t/2}^{t_0+\delta t/2} \left[\int_{\alpha_p} B\{\alpha_p - \alpha_s(t)\}\, I(\alpha_p)\, e^{-i2\pi\alpha_p u_p(t)}\, d\alpha_p \right] dt, \quad (C.1)$$

where δt is the integration time and $\hat{u}_p$ and $\hat{\alpha}_s$ are the mean spatial frequency and direction cosine, defined as

$$\hat{u}_p \equiv \frac{1}{\delta t}\int_{t_0-\delta t/2}^{t_0+\delta t/2} u_p(t)\, dt \quad \text{and} \quad \hat{\alpha}_s \equiv \frac{1}{\delta t}\int_{t_0-\delta t/2}^{t_0+\delta t/2} \alpha_s(t)\, dt. \quad (C.2)$$

In this section, we analyze the consequences of the antenna slewing on the accuracy of the wide-field synthesis.

C.1. Time averaging

In all interferometric observing modes, it is usual to adjust the integration time so that $u_p(t)$ can be approximated as $\hat{u}_p$. To do this, it is enough to ensure that $u_p(t)$ always varies less than the uv distance associated with tolerable aliasing (d_{alias}, see Sect. 4.2) during the integration time (δt)

$$\delta t \ll \frac{d_{\text{alias}}}{d_{\text{max}} \omega_{\text{earth}}} \quad \text{or} \quad \frac{dt}{1\,\text{sec}} \ll \frac{6900}{\theta_{\text{alias}}/\theta_{\text{syn}}}, \quad (C.3)$$

Table C.1. Definition of the symbols used to explore the influence of on-the-fly scanning on the measurement equation.

Symbol & Definition	
δt	Integration time
$\hat{\alpha}_s$	Scanned angle averaged during δt
$\hat{u}_p$	Spatial frequency averaged during δt
$\delta\alpha_s$	Angular distance scanned during δt
v_{slew}	Slew angular velocity of the telescope
A	Primary beam apodizing function
B_{eff}	Effective primary beam resulting from OTF scanning: $B_{\text{eff}}(\alpha) = \{B \star A\}(\alpha)$
ω_{earth}	Angular velocity of a spatial frequency due to Earth rotation

where d_{max} is the maximum baseline length, ω_{earth} is the angular velocity of a spatial frequency due to the Earth rotation (7.27×10^{-5} rad s^{-1}), θ_{alias} and θ_{syn} are respectively the minimum field of view giving a tolerable aliasing and the synthesized beam angular values.

C.2. Effective primary beam

Assuming that condition C.3 is ensured, we can write Eq. C.1 with the same form as Eq. 1 by the introduction of an effective primary beam (B_{eff}); *i.e.*,

$$V(\hat{u}_p, \hat{\alpha}_s) = \int_{\alpha_p} B_{\text{eff}}(\alpha_p - \hat{\alpha}_s)\, I(\alpha_p)\, e^{-i2\pi\alpha_p \hat{u}_p}\, d\alpha_p, \quad (C.4)$$

where $B_{\text{eff}}(\alpha_p - \hat{\alpha}_s) \equiv \frac{1}{\delta t}\int_{t_0-\delta t/2}^{t_0+\delta t/2} B\{\alpha_p - \alpha_s(t)\}\, dt. \quad (C.5)$

Using the following change of variables

$$\beta \equiv \alpha_s(t) - \hat{\alpha}_s, \quad d\beta = \frac{d\alpha_s(t)}{dt}dt \quad \text{or} \quad dt = \frac{d\beta}{v_{\text{slew}}(\beta)}, \quad (C.6)$$

we derive

$$B_{\text{eff}}(\alpha_p - \hat{\alpha}_s) = \int_\beta B\{(\alpha_p - \hat{\alpha}_s) - \beta\}\, A(\beta)\, d\beta \quad (C.7)$$

with $\quad A(\beta) \equiv \frac{1}{v_{\text{slew}}(\beta)\, \delta t} \Pi\left(\frac{\beta}{\delta\alpha_s}\right) \quad (C.8)$

and $\quad \delta\alpha_s \equiv \int_{t_0-\delta t/2}^{t_0+\delta t/2} v_{\text{slew}}(t)\, dt. \quad (C.9)$

In these equations, $v_{\text{slew}}(\beta)$ is the slew angular velocity of the telescope as a function of the sky position, $\delta\alpha_s$ is the angular distance covered during δt, A is an apodizing function, and $\Pi(\beta)$ is the usual rectangle function, which reproduces the finite character of the time integration.

C.3. Interpretation

The form of the measurement equation is conserved when averaging the visibility function over a finite integration time, as long as the true primary beam is replaced by an effective primary beam, which is the convolution of the true primary beam by an apodizing function. To go further, it is important to return to the two dimensional case. Indeed,
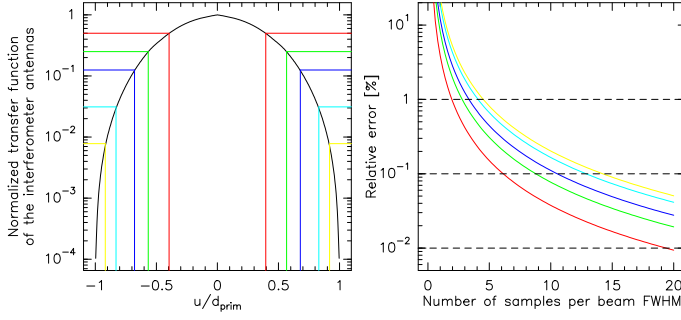
Fig. C.1. Assessement of the relative error implied by the use of the true primary beam instead of the effective primary beam when analyzing interferometric on-the-fly data sets. *Left:* Inverse Fourier transform of interferometer primary beam, $\overline{B}$ (*i.e.* the autocorrelation of the antenna illumination). *Right:* Relative error as a function of sampling rate of the primary beam. The curves of different colors show the results at different normalized uv distances (u/d_{prim}) from the center of $\overline{B}$.

the convolution must be done along the slewing direction, resulting in an effective primary beam elongated in a particular direction.

In principle, the equations derived in Sect. 3 can be accommodated just by replacing the true primary beam by its effective associate. In practice, the probability to take into account the effective primary beam is low because its shape varies with time. Indeed, it is often assumed that the sky is slewed along a straight line at constant angular velocity. Even in this simplest case, it is advisable to slew along at least two perpendicular directions to average systematic errors, implying two different effective primary beams. However, practical reasons may/will lead to complex scanning patterns: 1) The limitation of the acceleration when trying to image a square region leads to spiral or Lissajous scanning patterns; 2) The probable absence of derotators in future multi-beam receivers (B. Lazareff, private communication) implies the need to take into account the Earth rotation in the scanning patterns of the off-axis pixels.

C.4. Approximation accuracy

In the following, we thus ask what is the trade-off accuracy of using the true primary beam instead of the effective primary beam. The first point to mention is that using different scanning patterns somehow helps because the averaging process then makes the bias less systematic. Following Holdaway & Foster (1994), we quantify the accuracy lost in the Fourier plane. Indeed, the Ekers & Rots scheme tries to estimate missing sky brightness Fourier components from their measurements apodized by the Fourier transform of the primary beam. In the Fourier space, the above convolution just translates into a product. The Fourier transform of the apodizing function thus degrades the sensitivity of the measured visibility, $V(u_p, \alpha_s)$, to spatial frequencies at the edges of the interval $[u_p - d_{\mathrm{prim}}, u_p + d_{\mathrm{prim}}]$. To guide us in our quantification of the accuracy lost, we now explore the simplest case of linear scanning at constant velocity, where $v_{\mathrm{slew}}(\beta)$ is constant and $\delta\alpha_s = v_{\mathrm{slew}}\,\delta t$. The Fourier transform of the apodizing function is then a sinc function:

$$\overline{A}(u) = \mathrm{sinc}(u\,\delta\alpha_s). \tag{C.10}$$

The relative error implied by the use of the true primary beam instead of the effective primary beam is then

$$\frac{\overline{B}_{\mathrm{eff}}(u) - \overline{B}(u)}{\overline{B}_{\mathrm{eff}}(u)} = 1 - \frac{1}{\overline{A}(u)} = 1 - \frac{1}{\mathrm{sinc}(u\,\delta\alpha_s)}. \tag{C.11}$$

Fig. C.1 shows this relative error as a function of the number of samples per primary beam FWHM in the image plane (*i.e.*, $\theta_{\mathrm{fwhm}}/\delta\alpha_s$) for different uv distances (in units of d_{prim}). We see that we derive a 1% accuracy at all u when we sample the image plane at a rate of 5 dumps per primary beam. However getting a 0.1% accuracy needs quite high sampling rates (about 15). This must be compared with the accuracy of knowledge of $\overline{B}$.

We note that if a better accuracy is needed than the one achievable with the highest sampling rate, it is in theory possible to replace in the correlator software the rectangle apodizing function by another function which falls more smoothly. To avoid the loss of sensitivity inherent to the use of such an apodizing function (by throwing away data at the edges of the time interval of integration), would require, for instance, to half-overlap the integration intervals. This would imply more book-keeping in the correlator software and some noise correlation between the measured visibilities.